N69-40130
NASA CR-106347

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

Space Programs Summary 37-57, Vol. III

Supporting Research and Advanced Development

For the Period April 1 to May 31, 1969

CASE FILE
COPY

JET PROPULSION LABORATORY
CALIFORNIA INSTITUTE OF TECHNOLOGY
PASADENA, CALIFORNIA

June 30, 1969

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

Space Programs Summary 37-57, Vol. III

Supporting Research and Advanced Development

For the Period April 1 to May 31, 1969

JET PROPULSION LABORATORY
CALIFORNIA INSTITUTE OF TECHNOLOGY
PASADENA, CALIFORNIA

June 30, 1969

SPACE PROGRAMS SUMMARY 37-57, VOL. III

Copyright © 1969
Jet Propulsion Laboratory
California Institute of Technology

Prepared Under Contract No. NAS 7-100
National Aeronautics and Space Administration

Preface

The Space Programs Summary is a multivolume, bimonthly publication that presents a review of technical information resulting from current engineering and scientific work performed, or managed, by the Jet Propulsion Laboratory for the National Aeronautics and Space Administration. The Space Programs Summary is currently composed of four volumes:

Vol. I. *Flight Projects* (Unclassified)

Vol. II. *The Deep Space Network* (Unclassified)

Vol. III. *Supporting Research and Advanced Development* (Unclassified)

Vol. IV. *Flight Projects and Supporting Research and Advanced Development* (Confidential)

Contents

ADVANCED STUDIES

I. Future Projects 1

A. Lunar Surface Gravity Investigations

R. G. Brereton, NASA Code 195-06-10-01 1

SYSTEMS DIVISION

II. Systems Analysis Research 7

A. Equations of Motion of the General Three-Body Problem

R. Broucke and H. Lass, NASA Code 129-04-01-02 7

B. A Critique of the Scalar-Tensor Theory of Relativity

H. Lass, NASA Code 129-04-01-01 11

III. Computation and Analysis 14

A. The Design of an Algorithm for Solving Large Linear Least-Squares Systems Using Two Backup Tapes

R. J. Hanson, NASA Code 129-04-04-01 14

SPACE SCIENCES DIVISION

IV. Space Instruments 16

A. Electronics System for a Three-Axis High-Field Vector Helium Magnetometer

L. L. Lewyn, NASA Code 188-36-01-04 16

V. Lunar and Planetary Sciences 25

A. Error Analysis of the Linearization Method in Radiative Transfer Theory

K. D. Abhyankar and A. L. Fymat, NASA Code 185-47-33-01 25

B. Effect of Linearization of the Transfer Problem on the Computations of the Emergent Radiation Field

A. L. Fymat and K. D. Abhyankar, NASA Code 185-47-33-01 28

C. Lunar Mascons: Refined Calculation of the Model Gravity Profile for Mare Serenitatis

J. E. Conel, G. B. Holstrom, and E. Abbott, NASA Code 195-42-12-01 . . . 30

D. Retention and Release of Ion-Implanted Argon

H. C. Lord, III, NASA Code 195-42-12-01 33

Erratum 35

VI. Physics 36

A. Relativistic Many-Body Gravity Theory and Celestial Mechanics

F. B. Estabrook, NASA Code 129-02-07-02 36

Contents (contd)

TELECOMMUNICATIONS DIVISON

VII. Spacecraft Telemetry and Command 38

A. On the Capacity of a Cascade of Identical Discrete Memoryless Nonsingular Channels

M. K. Simon, NASA Code 186-68-04-28 38

B. Efficient Spacecraft Data Systems: Modifications to the *Mariner* Commutator to Adapt to the TOPS Mission

R. A. Easton and E. E. Hilbert, NASA Code 125-23-02-13 41

VIII. Spacecraft Telecommunications Systems 47

A. Frequency Tracking in an MFSK Receiver

H. D. Chadwick, NASA Code 150-22-17-04 47

B. Interference-to-Signal Ratios in a Frequency-Multiplexed PM Communication System

M. A. Koerner, NASA Code 186-68-04-11 54

C. Interference-to-Signal Ratios in a PM Single-Sideband Transmitter

M. A. Koerner, NASA Code 186-68-04-11 60

GUIDANCE AND CONTROL DIVISION

IX. Spacecraft Power 67

A. Electric Propulsion Power Conditioning

E. N. Costogue, NASA Code 120-26-04-05 67

B. Required Gamma Ray Shielding for Science Instruments on Radioisotope-Powered Spacecraft

M. A. Dore, NASA Code 120-27-06-01 68

C. Nuclear Power Systems Definition Studies

P. J. Gingo, NASA Code 120-27-06-01 73

D. Long-Term Life Test of the Spare *Mariner* Venus 67 Power Subsystem Hardware

A. Krug, NASA Code 120-33-08-11 76

E. Power Conditioner for a Thermionic Diode Array

D. J. Hopper, NASA Code 120-27-05-01 77

F. TWT Power Conditioning for Planetary Missions

D. J. Hopper, NASA Code 120-27-06-24 77

G. Power Conversion Circuit Development

D. J. Hopper, NASA Code 120-33-08-04 80

H. Solar Power System Definition Studies

H. Wick, NASA Code 120-33-05-01 81

I. Liquid Amalgam Electrodes

G. L. Juvinall, NASA Code 120-34-01-17 88

Contents (contd)

J. State-of-Charge Indicator Investigation
S. J. Krause, NASA Code 120-34-01-16 89

X. Spacecraft Control 93

A. Spurious Canopus Tracker Outputs
R. H. Burns, NASA Code 125-17-02-02 93

B. Nozzle Thrust at Leakage Flow Rates
F. G. Roselli-Lorenzini, NASA Code 731-13-01-03 95

C. Dynamic Simulation Utilizing Thrust Profile
L. L. Schumacher, NASA Code 160-76-01-01 97

D. Design and Prototype Development of the Pulse-Width Modulator and the Armature Current Direction Sensor for the Minimum-Energy Controller
Y. E. Sahinkaya, NASA Code 125-19-03-02 99

E. Digital Sun Sensor
L. F. Schmidt, NASA Code 186-68-02-19 103

F. On-Line Estimation of Atmospheric Parameters by an Approximate Nonlinear Filtering Technique
A. K. Bejczy, NASA Code 125-19-03-02 108

G. Solar Electric Spacecraft Attitude Control With Flexible Solar Arrays
G. E. Fleischer, NASA Code 125-19-04-01 114

XI. Guidance and Control Research 118

A. Guard-Ring Thermionic Diode Experiments
K. Shimada and P. L. Cassell, NASA Code 129-02-01-07 118

B. Thermal Noise of Injected Carriers
M-A. Nicolet, NASA Code 129-02-05-09 122

ENGINEERING MECHANICS DIVISION

XII. Electronics Parts Engineering 127

A. Power Pulse Method of Screening Resistors
K. C. Evans, NASA Code 186-70-01-04 127

ENVIRONMENTAL SCIENCES DIVISION

XIII. Space Simulators and Facility Engineering 130

A. Cryogenic Oscillators for Quartz Crystal Microbalances
E. Laue and C. H. Goldsmith, NASA Code 124-09-19-07 130

Contents (contd)

PROPULSION DIVISION

XIV. Solid Propellant Engineering 133

A. Heat Transfer in a Solid Propellant Grain During Ignition With Radiation Penetration and Convection Boundary Conditions
R. L. Klaus, NASA Code 128-32-50-01 133

B. Blast Wave Propagation in Air
O. K. Heiney, NASA Code 128-32-50-02 139

C. Pintle Nozzle Thrust Vector Control
L. D. Strand, NASA Code 128-32-52-01 144

XV. Polymer Research 148

A. Hookean, Newtonian, and Maxwellian Fillers in Hookean and Standard-Solid Matrices
J. Glucklich, NASA Code 128-32-43-01 148

B. Glassy-State Properties of Polymers: Epoxy Resins
E. F. Cuddihy and J. Moacanin, NASA Code 129-03-11-04 152

C. Simulation of the Venus Atmosphere
R. H. Silver, S. H. Kalfayan, and V. W. Bellin, NASA Code 186-58-13-09 . . . 157

D. Effects of Simulated Venus Atmosphere on Polymeric Materials
S. H. Kalfayan, R. H. Silver, and V. W. Bellin, NASA Code 186-58-13-09 . . . 160

E. Poly (isobutylene) Prepolymers
J. Miller, NASA Code 128-32-43-02 163

F. The Aging of Styrene–Butadiene Rubber: A Spectral Investigation
R. Rakutis and S. H. Kalfayan, NASA Code 129-03-11-04 166

G. Creep of Elastomers Undergoing Aging
J. Moacanin, J. J. Aklonis, R. F. Fedors, and R. F. Landel, NASA Code 129-03-11-04 169

XVI. Research and Advanced Concepts 173

A. Initial Operation of the Thermionic Diode Kinetics Experiment
H. G. Gronroos, M. L. Peelgren, and J. P. Davis, NASA Code 120-27-05-01 . . 173

B. Liquid-Metal MHD Power Conversion: Impinging Nozzle Experiment
L. G. Hays, NASA Code 120-27-06-03 177

C. Hollow-Cathode Operation in the SE-20C Ion Thruster
T. D. Masek and J. R. Womack, NASA Code 120-26-08-01 180

D. Laminarization of a Turbulent Boundary Layer in a Nozzle With Wall Cooling
R. F. Cuffel, L. H. Back, and P. F. Massier, NASA Code 128-31-54-02 181

E. Reference Mission Specification for 300-kWe Thermionic Spacecraft Design Study
N. K. Simon, NASA Code 120-27-06-23 185

Contents (contd)

XVII. Liquid Propulsion 187

A. Advanced Combustion Device Development

R. W. Riebling, NASA Code 731-12-04-04 187

B. Preliminary Burner Tests With Oxygen Difluoride and Diborane

F. Gerbracht, R. Scott, and R. Kushida, NASA Code 128-31-51-02 192

C. Porosity Surface Change of Shell 405 Catalyst, Preliminary Investigation

A. Karbin, NASA Code 731-13-01-07 194

XVIII. Propulsion Systems Analysis and Advanced Engineering

A. Sterilization and Impact Testing of Experimental Pyrotechnic Electronic Unit

M. L. Moore, NASA Code 186-58-08-01 200

I. Future Projects

ADVANCED STUDIES

A. Lunar Surface Gravity Investigations,

R. G. Brereton

1. Introduction

The surface gravity of the moon is only one-sixth that of the earth; therefore, gravity anomalies on the moon resulting from a density contrast in lunar material will comprise a larger part of the total-field measurement than similar measurements on earth. Lunar gravity anomalies may be caused by local near-surface density contrasts in rock units, as between the regolith and basement rocks, or perhaps by regional isostatic phenomena where the moon's crust is out of isostatic equilibrium because of ancient frozen tidal effects or crustal overloading by ejecta from large meteor impacts. The Carpathians and Apennines are possible examples of crustal overloading from the Imbrium impact event; also, the evidence for so-called mascons in some of the circular maria presents an interesting example of lunar gravity problems. Surface gravity data from profiles across virtually all lunar structures and contacts are desirable since these data may be critical to an understanding of the origin and evolution of these features and even the moon itself.

Two modes of lunar surface gravity investigations are probable—viz., fixed station operation and gravity exploration with an automated roving vehicle. Both modes of operation may be required in order to provide the quality and quantity of observations for the solution of lunar gravity problems. Gravity observations on the moon will have to be corrected for both a "free-air" and "Bouguer" effect, and under more limited circumstances, a topographic correction and a tidal correction may be required.

The best gravity instrument for early lunar missions is probably a conventional spring-mass gravimeter in contrast to a torsion balance or pendulum. The state-of-the-art for design of this instrument is highly advanced, since it is the primary instrument for terrestrial gravity prospecting. Terrestrial gravimeters are required to detect changes in gravity of the order of 1 part/10^8.

2. The Lunar Gravity Field

The lunar gravity field intensity and its anomalous components can be defined by potential theory. Consider a gravity field of force and let us insert a suitable test body into this field. There will be a force exerted by the field on the test body which we can call F. We now carry this test body from its original position around an arbitrary closed path. The work done by the field on the body as the latter undergoes a small displacement, say ds, is given by

$$dw = \mathbf{F}_s \, ds$$

where $\mathbf{F}_s$ is the component of F in the direction ds. If the field is conservative, we have for the condition of any closed path:

$$\oint \mathbf{F}_s \, ds = 0$$

When the above condition is realized, we can say that a potential function exists defined between any two points in space, say

$$\sqcup_B - \sqcup_A = \int_A^B \mathbf{F}_s \, ds$$

where $\sqcup$ represents the potential at a designated position in space. It is usual to set the potential energy of the test particle equal to zero at points infinitely distant from the region of space where the gravity field exists. If we do this, the potential at a point P of the field becomes

$$\sqcup_P = \int_\infty^P \mathbf{F}_s \, ds$$

The above equation can be shown to be the same as

$$\sqcup_P = \int_\infty^P F \cdot d\mathbf{r}$$

where $\mathbf{r}(x, y, z)$ is the vector from the origin to the test particle. The potential is seen to be a function of the coordinates x, y, z of P, and by differentiation we find that

$$d\sqcup = F \cdot d\mathbf{r}$$

We can define

$$d\sqcup = \mathbf{i}\frac{\partial U}{\partial x} + \mathbf{j}\frac{\partial U}{\partial y} + \mathbf{k}\frac{\partial U}{\partial z} \cdot d\mathbf{r}$$

where the quantity

$$\mathbf{i}\frac{U}{x} + \mathbf{j}\frac{U}{y} + \mathbf{k}\frac{U}{z} = \text{grad}\sqcup \equiv \nabla\sqcup$$

Then

$$F = \nabla\sqcup$$

In general, where $\sqcup$ represents the potential function for a spherical volume of matter as the moon,

$$\sqcup = -G\int_v \rho_v \frac{dv}{\mathbf{r}}$$

where G is the gravitational constant, ρ_v is the density of a unit volume of matter, and $\mathbf{r}$ is the magnitude of the radius vector from the elemental mass to the point in question. Then the field intensity is represented by the relationship

$$F = \nabla\sqcup$$

where ∇ is a vector operator of the form

$$\mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z}$$

For points outside a spherical body of mass M and radius R, as the moon, the magnitude of the general field intensity is

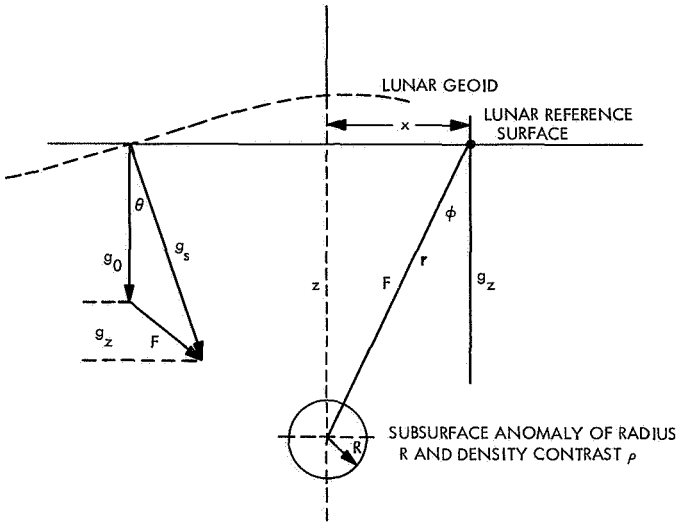
$$F = G\frac{M}{R^2}$$

Potential functions can be derived for a variety of geometrical forms and mass distributions, and these equations can be quite useful for defining the gravity field intensity for these bodies.

Since the gravitational constant and the mass of the moon are both known to several significant figures, the theoretical value of the gravity field of the moon can be calculated as a function of the lunar radius. If we assume the diameter of the moon is 3476 km and its mass is 7.35×10^{25} g, then the gravitational field on its surface is approximately 163 cm/s^2 (163 gals). Errors in our knowledge of the lunar radius may affect this value by as much as ± 100 milligals and crustal anomaly by a like amount; but in general, if the shape of the lunar geoid is known and the value of the acceleration of gravity is known at one point, then the acceleration of lunar gravity can be computed for any other point. And conversely, if the value of the acceleration of gravity is known at all points on the lunar geoid, then the form of the level surface can be deduced.

Anomalies in the gravity field of the moon can be attributed to density contrasts in surface and subsurface rocks, so-called terrain effects which lead to the free-air, Bouguer, and topographic corrections, and finally, shape effects.

The most practical means of representing the gravity effects resulting from density contrasts in rocks is to calculate the vertical component of gravity g_z along a profile over the center of the body. It should be realized that field gravity meters are generally designed to respond to this particular component. Figure 1 illustrates the relationship of the various components of attraction for a buried sphere in the lunar crust.



g_0 = THEORETICAL GRAVITY ON LUNAR REFERENCE SURFACE EQUAL TO APPROXIMATELY G M/R^2

g_s = GRAVITY AT STATION FROM EFFECT OF g_0 AND ANOMALY

g_z = $g_s - g_0$ REPRESENTS THE GRAVITATIONAL ANOMALY THAT WOULD BE REGISTERED BY A CONVENTIONAL GRAVIMETER WHERE θ IS SMALL

Fig. 1. Components of the lunar gravity field

In cases where the anomalous body has a shape that can be expressed mathematically, then the anomalous component of gravity can be computed by the methods of potential theory using summation techniques for line, surface, and volume integrals for such forms as spheres, cylinders, flat disks, cones, slabs, etc.; also, expressions can be derived even when the form presents an acute or obtuse angle with the surface, but the mathematical expression for other than the regular geometrical form which is oriented horizontally or vertically to the surface can be quite complicated and difficult to derive. Under these circumstances, g_z is usually determined by the aid of various graphical and mechanical computation techniques.

It must be realized that the interpretation of gravity anomalies in terms of real geological structures requires external control. It is impossible to calculate directly from a gravity effect a unique mass distribution, shape, depth, size, or density for an anomalous body. The degree of fit between computed and observed effects is dependent on the closeness and precision of the data, on the uniqueness of the data, and primarily on the amount of geological control that can be exercised to limit the inherent ambiguity and lend uniqueness to the assumed mass distribution. For example, from a single gravity profile across say a sphere, a horizontal cylinder, and a vertical sheet, it would likely be hard to tell these forms apart; however, wider areal coverage would show the elongated dimensions of the cylinder and the vertical sheet, and if the vertical sheet were quite thick, then closely spaced precise data may help to distinguish it from the cylinder. Also, the addition of geological information in regard to the depth, size, shape, density, etc., of the anomalous mass would provide the control for interpretation.

Gravity measurements on the lunar surface will have to be corrected for so-called terrain and topographic effects. This is necessary in order to relate all observations to a common datum. If the moon were a perfect fluid with no lateral variations in density, its surface would correspond to a spheroid and the direction of gravity over this level surface would be everywhere perpendicular to it. Obviously, the lunar surface is not always parallel with a spheroid of reference, and gravity measurements made on the actual surface will have to be corrected to the reference surface if gravity values are to be used for defining the shape of the moon or crustal structures.

The important terrain corrections are the free-air and Bouguer. The free-air correction takes into account the fact that gravity varies with elevation and measurements made at a higher elevation are farther from the center of mass and, therefore, have a smaller gravity value than measurements at a lower elevation. Since the moon's mass can be considered to be concentrated at its center, then if g_0 represents the value of gravity at its surface and g_h the value at height h, the inverse-square law tells us that the gravity difference between the two points is

$$\frac{g_0}{g_h} = \frac{(R+h)^2}{R^2} = \frac{R^2 + 2Rh + h^2}{R^2}$$

and

$$\frac{g_0}{g_h} = 1 + \frac{2h}{R} + \frac{h^2}{R^2}$$

Because h will be small compared to R, the term h^2/R^2 can be neglected, and we can write

$$g_0 - g_h = \frac{2hg_h}{R}$$

This is also the same as the vertical gradient of gravity,

$$\frac{\partial g}{\partial r} = \frac{2g}{r}$$

and on the moon its value is

0.187 milligals/m

The lunar free-air correction adjusts gravity observations for points made at different elevations to a common datum, but it does not account for the change in mass under the station in order to change its elevation. In other words, to change the elevation of surface observation points requires that a specific volume of surface rock material be either added to or subtracted from the datum. This effect is called the Bouguer correction. It is always opposite in sign to the free-air correction. It is usual to assume either a slab or cylinder of infinite horizontal dimensions and finite vertical dimensions in order to compute an expression for the Bouguer correction. The value for the Bouguer correction is

$$\Delta B = 2\pi G\rho h = 0.042\rho \text{ milligals/m}$$

If we assume an average density for lunar crustal rocks of say 2.67 g/cm^3, then the Bouguer correction is

$$0.112 \text{ milligals/m}$$

and the combined lunar terrain correction, that is, the free-air minus the Bouguer, is approximately 0.075 milligals/m. Interestingly, this value is less than one-half of the correction required on earth, and this has an important effect on field operations on the moon since it reduces the accuracy for elevation control. It should be realized that the magnitude of the Bouguer correction will change for sedimentary areas in contrast to igneous areas, and also in response to regional investigations in contrast to local studies. The Bouguer correction is very sensitive to rock densities, and in a local survey in hilly country it can serve as a precise tool for determining the average density of local rock units.

A topographic correction will be required only in hilly areas. It may be thought of as a supplemental correction to the smooth infinite slab hypothesized in making the Bouguer correction, for it accounts for all material projecting above the Bouguer plain and all material needed to fill in hollows below the plain. It is always positive, and for most survey work, it will not be troublesome.

Tidal corrections for gravity measurements on the lunar surface, caused by the differential attraction between perigee and apogee, may be larger than expected if the crust of the moon yields to the external force. The external force difference between perigee and apogee is very small, of the order of a milligal, but if the lunar crust yields to this force, then gravity readings on the surface will be affected. It is not likely that this effect will be significant for mobile surface exploration; however, it may have some important scientific repercussions on fixed station gravity observations.

3. Gravity Experiments and Objectives

Lunar surface gravity observations will need to include both fixed-station and exploration-type measurements. Fixed-station gravity experiments will provide fundamental data towards the task of lunar geodesy, that is, determining the size and shape of the moon, providing control points for mapping and charting work, and providing a gravity base. In addition, these stations may provide some interesting scientific data on the strength of the lunar crust and the physical properties and nature of the lunar interior by observing the moon's crustal tides through several orbital cycles around the earth. The actual yielding of the solid earth to tidal forces is approximately a foot, but the tidal yielding on the solid moon from the pull of the earth between apogee and perigee may be several meters. This experiment needs to be carried out either at the limb or center of the lunar disk where tidal effects are maximum. An accurate ranging experiment would be a most desirable associate experiment.

A roving vehicle capability for the moon adds a whole new dimension to our task of lunar gravity observations. Specifically, traverse-type gravity observations can

(1) Provide information on lunar isostasy. Are the mountain masses ringing the Imbrium, Humorium, and other circular maria uncompensated masses supported by a strong lunar crust? What is the strength of the lunar crust and the rate of adjustment for events ranging in age from the Imbrium through the Orientale? Was the adjustment rate different for different ages and places?

(2) Provide information on the lunar crust. What is the depth of the lunar crust? Are the maria and terra surface images of deep crustal phenomena? What is the density relationship between major crustal units?

(3) Provide information on the shape of the moon and in support of lunar geodetic data. Does the moon have a frozen tidal bulge? Is the front side of the moon, as indicated by tracking data, slightly depressed?

(4) Provide information on the anomalous gravity field for such lunar features as impact craters, volcanic forms, mascons, ridges, faults, rills, basins, etc. How

does the gravity field for these features compare with terrestrial analogs? What is the density of lunar surface and near-subsurface material? Can lunar stratigraphic units be mapped by gravity survey techniques?

For determining the gravity effect for large structural units, the concept of a free-air and Bouguer anomaly are quite useful. For example, the free-air anomaly can be defined as

g_f = observed gravity + free air correction − theoretical gravity

The observation of the free-air effect on a traverse over a lunar mountain mass, say as the Apennines, would offer a direct test of isostasy. Thus, a free-air anomaly near zero would indicate perfect compensation, or that the mountain mass was hollow. Such a traverse, taking into account a Bouguer effect, could indicate the degree of crustal compensation for the mass of the mountain and also provide data on the density of the mountain and lunar crust.

The Bouguer anomaly can be defined as

g_B = observed gravity + free-air correction − Bouguer correction + topographic correction − theoretical gravity

A Bouguer anomaly near zero for a mountain area signifies a lack of compensation or little isostatic adjustment. In general, for an uncompensated mountain on the moon, one would expect a Bouguer anomaly of less than a few milligals and a free-air anomaly of say 100 milligals, while if the moon were plastic for either a particular period or location, then the Bouguer anomaly for a surface feature would be large and strongly negative and the free-air anomaly quite small.

Lunar gravity observations may provide an insight into the interpretation of terrestrial data and conversely, terrestrial analogs will be useful for understanding lunar data. For example, a terrestrial volcanic feature usually gives a positive anomaly, while a terrestrial impact crater, due to the low-density brecciated fill within the crater, usually gives a negative anomaly. Also, often the raised rim of an impact feature will give a small positive anomaly due to the dense material that has been forced-up around the crater. It will be mutually instructive to compare field observations for these features on the moon and on the earth.

Lunar gravity observations may be directed to a search for the consequences, or proof, if you wish, of a particular hypothesis—for example, the lunar crust evolved through large scale melting and differentiation at a particular period in geologic time, or the lunar crust has only selectively melted, or the lunar crust has a so-called "raisin bread" structure. Gravity survey data could be directed to a search for the consequences of these hypotheses, and most probably, gravity data would provide the most discriminatory information for a solution.

Gravity observations from the lunar surface can be expected to provide unique data that will be applicable towards the solution of a host of lunar geological and geophysical problems.

4. Instruments and Operating Procedures

Gravity meters for terrestrial use are designed to read with a precision of at least 10^{-8} of the total field, that is, to a precision of 0.01 milligal. More sensitive meters are possible and have been built for special purposes; however, the limit of 0.01 milligal is set in normal field use by the noise level of small topographic changes and practical field procedures. For work on the moon, there is no reason why this basic instrument precision should not be retained; however, the total lunar field is only one-sixth that of the earth's and this, in conjunction with more difficult lunar operating procedures, a greater range of ambient temperature, and the requirement for extended periods of unattended operation, will result in a much lower applied accuracy. For fixed station observations, an accuracy of perhaps 10^{-6} of the total lunar field is certainly practical, but measurements from an automated roving vehicle, because of uncertainties in the magnitude of the terrain effect, leveling effects, and other small cumulative effects, will probably be reduced in accuracy to ± 3 milligals.

With the lunar terrain effect averaging 0.075 milligal/m, it will be desirable to know the elevation of the rover and the elevation of the terrain around it in relation to a fixed datum reference, or at least a series of such references, to an accuracy of ± 10 m in order to control the terrain error to approximately one-fourth of the desired survey accuracy. Since density contrasts in lunar rocks, that is, the cause of gravity anomalies, are expected to comprise a larger part of the total gravity field measurement on the moon than similar measurements on earth would, it may be practical to reduce the overall survey accuracy to ± 5 milligals, thus reducing topographic control for the terrain correction. If error effects from

other sources such as drift, calibration, etc., appear not to be so important, then it may be possible to relax the requirement for topographic control and still retain the desired survey accuracy of ±3 milligals.

The most practical and adaptable terrestrial gravity meter for lunar use would be an astatic type similar in design to the well known WORDEN[1] gravimeter. In this device, variations in gravity are measured by observing the changes in position of a mass supported by a gravity spring system. The state-of-the-art in design and operation of this instrument is very high and it could be readily modified for lunar use. It must be realized that this type of instrument is not designed to give the absolute value of gravity, but rather the change in relative gravity. This means it must be calibrated in terms of an absolute standard, and for the range of gravity values that will likely be encountered. Herein lies our biggest source of error for using these instruments at fixed sites to determine the absolute gravity, because each instrument at each site will have a slightly different calibration and response. For rover surveys, except where it becomes desirable to tie surveys together or relate a survey to a fixed site, this will not present an error greater than the ±3 milligals desired. The sources of instrument error for lunar operations are:

(1) *Calibration.* This can probably be accomplished to an accuracy of ±2 milligals. For the rover case where the instrument must have a range of response of 500 milligals, an additional error of about ½ of 1% may be encountered at the extremes of this range.

[1]Designed by Texas Instruments, Incorporated.

(2) *Drift.* This is caused by changes in the elastic or frictional properties of the gravity meter. For quartz-fiber-type instruments, this amounts to only a few tenths of a milligal per month. A drift of less than ±0.5 milligal per month is achievable in instruments designed for lunar rovers.

(3) *Leveling.* An angular error of 15 min of arc will produce an error of approximately 1.6 milligals on the moon; therefore, leveling of the rover gravity instrument before reading will be an important step in the operating procedure. Because of oscillations in the rover, it will probably be necessary to lower the gravimeter to the lunar surface when measurements are desired.

(4) *Thermal control.* Quartz-fiber-type meters used on earth have been compensated for the normal terrestrial temperature range, so corrections for this are not normally required. On the moon, the greater range of temperature will require a different compensation technique, or perhaps active thermal control for fixed station observations. Good insulation, along with achievable compensation, will probably keep rover instrument errors to less than ±0.5 milligals per lunation.

The spacing of surface gravity measurements will be a function of the size of the feature being observed. For the determination of the density of surface material, a close-space profile over several low hills and valleys would be required whereas a traverse to investigate a regional phenomenon may have observation points separated by several miles.

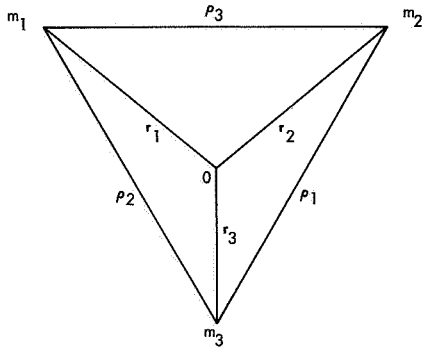
II. Systems Analysis Research

SYSTEMS DIVISION

A. Equations of Motion of the General Three-Body Problem, *R. Broucke and H. Lass*

The equations of motion for the general Newtonian three-body problem take on symmetric forms when one considers the three relative accelerations of the bodies. From these equations one quickly surmises the well-known classical properties of the three-body problem.

The derivation of these equations is obtained from the following considerations. Let the center of mass remain at the origin of the three-dimensional coordinate system and consider the vectors ρ_1, ρ_2, ρ_3 forming the closed triangle



Now we have

$$\left.\begin{aligned} \rho_1 &= \mathbf{r}_2 - \mathbf{r}_3 \\ \rho_2 &= \mathbf{r}_3 - \mathbf{r}_1 \\ \rho_3 &= \mathbf{r}_1 - \mathbf{r}_2 \end{aligned}\right\} \tag{1}$$

and

$$\rho_1 + \rho_2 + \rho_3 = 0$$

along with the fact that the center of mass remains at the origin, namely,

$$m_1\mathbf{r}_1 + m_2\mathbf{r}_2 + m_3\mathbf{r}_3 = 0 \tag{2}$$

One needs Eq. (2) in order that $\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3$ can be uniquely determined if any two of the vectors ρ_1, ρ_2, ρ_3 are specified. The system of Eqs. (1) and (2) has a unique solution since the determinant of the system is

$$\sum_{i=1}^{3} m_i > 0$$

One can easily show that

$$\left.\begin{aligned}\mathbf{r}_1 &= \frac{1}{\mu}(m_3\boldsymbol{\rho}_2 - m_2\boldsymbol{\rho}_3)\\ \mathbf{r}_2 &= \frac{1}{\mu}(m_1\boldsymbol{\rho}_3 - m_3\boldsymbol{\rho}_1)\\ \mathbf{r}_3 &= \frac{1}{\mu}(m_2\boldsymbol{\rho}_1 - m_1\boldsymbol{\rho}_2)\end{aligned}\right\} \tag{3}$$

where $\mu = m_1 + m_2 + m_3$.

The equations of motion of the three particles are

$$\ddot{\mathbf{r}}_1 = -m_2\frac{\mathbf{r}_1 - \mathbf{r}_2}{\rho_2^3} - m_3\frac{\mathbf{r}_1 - \mathbf{r}_3}{\rho_3^3}$$

$$\ddot{\mathbf{r}}_2 = -m_1\frac{\mathbf{r}_2 - \mathbf{r}_1}{\rho_1^3} - m_3\frac{\mathbf{r}_2 - \mathbf{r}_3}{\rho_3^3}$$

$$\ddot{\mathbf{r}}_3 = -m_1\frac{\mathbf{r}_3 - \mathbf{r}_1}{\rho_1^3} - m_2\frac{\mathbf{r}_3 - \mathbf{r}_2}{\rho_2^3}$$

From Eq. (1) and the equations of motion, we obtain

$$\begin{aligned}\ddot{\boldsymbol{\rho}}_1 &= \ddot{\mathbf{r}}_2 - \ddot{\mathbf{r}}_3\\ &= -\frac{m_3\boldsymbol{\rho}_1}{\rho_1^3} + \frac{m_1\boldsymbol{\rho}_3}{\rho_3^3} - \frac{m_2\boldsymbol{\rho}_1}{\rho_1^3} + \frac{m_1\boldsymbol{\rho}_2}{\rho_2^3}\\ &= -\frac{(m_1 + m_2 + m_3)}{\rho_1^3}\boldsymbol{\rho}_1 + m_1\left(\frac{\boldsymbol{\rho}_1}{\rho_1^3} + \frac{\boldsymbol{\rho}_2}{\rho_2^3} + \frac{\boldsymbol{\rho}_3}{\rho_3^3}\right)\end{aligned} \tag{4}$$

and two similar equations for $\boldsymbol{\rho}_2$ and $\boldsymbol{\rho}_3$.

We now let

$$m_0^2 = \frac{m_1m_2m_3}{\mu}$$

according to Deprit and Delie (Ref. 1). We also define the three quantities n_1, n_2, n_3, which are dimensionally similar to m_1, m_2, m_3:

$$n_1 = \frac{m_2m_3}{\mu}, \qquad n_2 = \frac{m_1m_3}{\mu}, \qquad n_3 = \frac{m_1m_2}{\mu}$$

The above definitions yield

$$\sum_{i=1}^{3}\frac{1}{n_i} = \frac{\mu}{m_0^2}$$

Multiplying Eq. (4) by n_1 will give us a new form for the equations of motion:

$$\left.\begin{aligned}n_1\frac{d^2\boldsymbol{\rho}_1}{dt^2} &= -\frac{\mu n_1}{\rho_1^3}\boldsymbol{\rho}_1 + m_0^2\left(\frac{\boldsymbol{\rho}_1}{\rho_1^3} + \frac{\boldsymbol{\rho}_2}{\rho_2^3} + \frac{\boldsymbol{\rho}_3}{\rho_3^3}\right)\\ n_2\frac{d^2\boldsymbol{\rho}_2}{dt^2} &= -\frac{\mu n_2}{\rho_2^3}\boldsymbol{\rho}_2 + m_0^2\left(\frac{\boldsymbol{\rho}_1}{\rho_1^3} + \frac{\boldsymbol{\rho}_2}{\rho_2^3} + \frac{\boldsymbol{\rho}_3}{\rho_3^3}\right)\\ n_3\frac{d^2\boldsymbol{\rho}_3}{dt^2} &= -\frac{\mu n_3}{\rho_3^3}\boldsymbol{\rho}_3 + m_0^2\left(\frac{\boldsymbol{\rho}_1}{\rho_1^3} + \frac{\boldsymbol{\rho}_2}{\rho_2^3} + \frac{\boldsymbol{\rho}_3}{\rho_3^3}\right)\end{aligned}\right\} \tag{5}$$

by considering, similarly, the acceleration terms for $\boldsymbol{\rho}_2, \boldsymbol{\rho}_3$.

Equation (5) can be looked upon as representing the equations of motion of three particles of masses $n_i, i = 1, 2, 3$, with position vectors $\boldsymbol{\rho}_i, i = 1, 2, 3$. Each particle is acted upon by a force of attraction due to a fixed mass μ, located at the origin, as well as by a perturbing force

$$m_0^2\left(\frac{\boldsymbol{\rho}_1}{\rho_1^3} + \frac{\boldsymbol{\rho}_2}{\rho_2^3} + \frac{\boldsymbol{\rho}_3}{\rho_3^3}\right)$$

which may be thought of as the force that three unit masses placed at the location of the masses n_i would exert on a mass m_0^2 placed at the origin.

To verify that Eq. (5) is consistent with $\boldsymbol{\rho}_1 + \boldsymbol{\rho}_2 + \boldsymbol{\rho}_3 \equiv 0$, we note that

$$\sum_{i=1}^{3}\frac{d^2\boldsymbol{\rho}_i}{dt^2} = -\mu\sum_{i=1}^{3}\frac{\boldsymbol{\rho}_i}{\rho_i^3} + m_0^2\sum_{i=1}^{3}\frac{1}{n_i}\sum_{i=1}^{3}\frac{\boldsymbol{\rho}_i}{\rho_i^3} \equiv 0 \tag{6}$$

since

$$m_0^2\sum_{i=1}^{3}\frac{1}{n_i} = \mu$$

The energy integral can be obtained from Eq. (5) by noting that

$$\begin{aligned}\sum_{i=1}^{3}n_i\frac{d\boldsymbol{\rho}_i}{dt}\cdot\frac{d^2\boldsymbol{\rho}_i}{dt^2} &= -\mu\sum_{i=1}^{3}\frac{n_i}{\rho_i^3}\boldsymbol{\rho}_i\cdot\frac{d\boldsymbol{\rho}_i}{dt}\\ &\quad + m_0^2\left(\frac{\boldsymbol{\rho}_1}{\rho_1^3} + \frac{\boldsymbol{\rho}_2}{\rho_2^3} + \frac{\boldsymbol{\rho}_3}{\rho_3^3}\right)\cdot\sum_{i=1}^{3}\frac{d\boldsymbol{\rho}_i}{dt}\\ &= -\mu\sum_{i=1}^{3}\frac{n_i}{\rho_i^3}\boldsymbol{\rho}_i\cdot\frac{d\boldsymbol{\rho}_i}{dt}\end{aligned} \tag{7}$$

since

$$\sum_{i=1}^{3} \frac{d\boldsymbol{\rho}_i}{dt} = 0$$

Thus,

$$\frac{d}{dt}\left[\frac{1}{2}\sum n_i \left(\frac{d\boldsymbol{\rho}_i}{dt}\right)^2\right] = \mu \frac{d}{dt}\left(\sum \frac{n_i}{\rho_i}\right) \tag{8}$$

which yields the energy integral

$$\frac{1}{2}\sum_{i=1}^{3} n_i \left(\frac{d\boldsymbol{\rho}_i}{dt}\right)^2 - \mu \sum_{i=1}^{3} \frac{n_i}{\rho_i} = E \tag{9}$$

The angular momentum integral is obtained by noting that

$$\begin{aligned}\sum_{i=1}^{3} n_i \boldsymbol{\rho}_i \times \frac{d^2\boldsymbol{\rho}_i}{dt^2} &= m_0^2\left[\frac{\boldsymbol{\rho}_1 \times \boldsymbol{\rho}_2}{\rho_2^3} + \frac{\boldsymbol{\rho}_1 \times \boldsymbol{\rho}_3}{\rho_3^3} + \frac{\boldsymbol{\rho}_2 \times \boldsymbol{\rho}_1}{\rho_1^3} + \frac{\boldsymbol{\rho}_2 \times \boldsymbol{\rho}_3}{\rho_3^3} + \frac{\boldsymbol{\rho}_3 \times \boldsymbol{\rho}_1}{\rho_1^3} + \frac{\boldsymbol{\rho}_3 \times \boldsymbol{\rho}_2}{\rho_2^3}\right] \\ &= m_0^2\left[\frac{(\boldsymbol{\rho}_2 + \boldsymbol{\rho}_3) \times \boldsymbol{\rho}_1}{\rho_1^3} + \frac{(\boldsymbol{\rho}_1 + \boldsymbol{\rho}_3) \times \boldsymbol{\rho}_2}{\rho_2^3} + \frac{(\boldsymbol{\rho}_1 + \boldsymbol{\rho}_2) \times \boldsymbol{\rho}_3}{\rho_3^3}\right] \\ &\equiv 0\end{aligned} \tag{10}$$

since $(\boldsymbol{\rho}_2 + \boldsymbol{\rho}_3) \times \boldsymbol{\rho}_1 = -\boldsymbol{\rho}_1 \times \boldsymbol{\rho}_1 = 0$, etc.

From

$$\boldsymbol{\rho}_i \times \frac{d^2\boldsymbol{\rho}_i}{dt^2} = \frac{d}{dt}\left(\boldsymbol{\rho}_i \times \frac{d\boldsymbol{\rho}_i}{dt}\right)$$

it follows that

$$\sum n_i \boldsymbol{\rho}_i \times \frac{d\boldsymbol{\rho}_i}{dt} = -\mathbf{J} \tag{11}$$

with $-\mathbf{J}$ the constant angular momentum.

In a similar way the well-known Lagrange identity may be obtained from our equations of motion (5). Let the moment of inertia of the three masses be defined by

$$I = \sum_{i=1}^{3} n_i (\boldsymbol{\rho} \cdot \boldsymbol{\rho})$$

Then, we can obtain a very simple expression for the second derivative of I. We have

$$\frac{dI}{dt} = 2\sum_{i=1}^{3} n_i \left(\boldsymbol{\rho}_i \cdot \frac{d\boldsymbol{\rho}_i}{dt}\right)$$

$$\begin{aligned}\frac{d^2I}{dt^2} &= 2\sum_{i=1}^{3} n_i \left[\boldsymbol{\rho}_i \cdot \frac{d^2\boldsymbol{\rho}_i}{dt^2} + \left(\frac{d\boldsymbol{\rho}_i}{dt^2}\right)^2\right] \\ &= 2\sum_{i=1}^{3} \boldsymbol{\rho}_i \cdot \left[-\mu \frac{n_i}{\rho_i^3}\boldsymbol{\rho}_i + m_0^2 \sum_{j=1}^{3} \frac{\boldsymbol{\rho}_j}{\rho_j^3}\right] + 4T \\ &= -2\mu \sum \frac{n_i}{\rho_i} + 4T \\ &= 2E - 2T + 4T = 2(E + T)\end{aligned}$$

where E is the energy constant as defined in Eq. (9) and T the kinetic energy.

An immediate consequence of Eq. (5) is the following result due to Lagrange, and called the equilateral triangle solution. Let $\rho_1 = \rho_2 = \rho_3 = \rho$ initially. Can the motion continue such that $\rho_i \equiv \rho$ throughout all of time, $i = 1, 2, 3$? Let the three bodies remain in a fixed plane. From

$$\sum_{i=1}^{3} \boldsymbol{\rho}_i \equiv 0$$

the equations of motion (5) uncouple, and become the familiar equations of Keplerian motion.

Thus, of necessity,

$$\left.\begin{aligned} &\rho^2\left(\frac{d\phi_i}{dt}\right) = k_i = \text{constant}, \qquad i = 1,2,3 \\ &\rho\left(\frac{d\phi_i}{dt}\right)^2 = \frac{\mu}{\rho^2} \end{aligned}\right\} \tag{12}$$

with ϕ_i the angle that $\boldsymbol{\rho}_i$ makes with a fixed line, say, the x-axis. Thus, all three particles of masses (n_1, n_2, n_3) must have an angular momentum $k_i^2 = \mu\rho$. Conversely, if Eq. (12) is satisfied, then $\rho_i \equiv \rho \equiv$ constant is a solution of Eq. (5). Since the motion is in a plane, we have $\dot{\boldsymbol{\rho}}_i = \boldsymbol{\omega} \cdot \boldsymbol{\rho}_i$, $\boldsymbol{\omega} \cdot \boldsymbol{\rho}_i = 0$, so that

$$\boldsymbol{\rho}_i \cdot \frac{d\boldsymbol{\rho}_i}{dt} = \rho^2\boldsymbol{\omega}$$

and Eq. (11) yields

$$\left(\sum_{i=1}^{3} n_i\right)\rho^2\boldsymbol{\omega} = \mathbf{J}$$

The solution considered above refers to circular orbits for each of the masses (m_1, m_2, m_3). Since the equations of motion (5) uncouple if $\rho_1 = \rho_2 = \rho_3 = \rho\,(k)$, it follows that the orbits can be similar conic sections with arbitrary eccentricity, yielding the equilateral triangle solution whose sides vary with time.

We return now to an alternative derivation of Eq. (5) using the theory of Lagrange multipliers. The Lagrangian is

$$L = \frac{1}{2}\sum_{i=1}^{3} m_i\dot{\mathbf{r}}_i^2 + \sum_{(i,j,k)}\frac{m_i m_j}{\rho_k} \tag{13}$$

with

$$\sum_{(i,j,k)}$$

a sum over the cyclic perturbations of i, j, k, i.e.,

$$\sum_{(i,j,k)}\frac{m_i m_j}{\rho_k} = \frac{m_1 m_2}{\rho_3} + \frac{m_2 m_3}{\rho_1} + \frac{m_3 m_1}{\rho_2}$$

Making use of Eq. (2) and $(\dot{\boldsymbol{\rho}}_1 + \dot{\boldsymbol{\rho}}_2 + \dot{\boldsymbol{\rho}}_3)^2 = 0$ yields

$$L = \sum_{i=1}^{3}\frac{n_i}{2}\left(\frac{\dot{\boldsymbol{\rho}}_i^2}{2} + \frac{\mu}{\rho_i}\right) \tag{14}$$

subject to the constraint

$$\sum_{i=1}^{3}\boldsymbol{\rho}_i = 0$$

If we let $\boldsymbol{\rho}_i = (X_i, Y_i, Z_i)$, then

$$\bar{L} = \sum\frac{n_i}{2}\left[\frac{(\dot{X}_i^2 + \dot{Y}_i^2 + \dot{Z}_i^2)}{2} + \frac{\mu}{\rho_i}\right] + \alpha\sum_{i=1}^{3} X_i + \beta\sum_{i=1}^{3} Y_i + \gamma\sum_{i=1}^{3} Z_i \tag{15}$$

satisfies the Euler–Lagrange equations, with α, β, γ Lagrange multipliers (not necessarily constants). From

$$\frac{d}{dt}\left(\frac{\partial\bar{L}}{\partial\dot{X}_i}\right) - \frac{\partial\bar{L}}{\partial X_i} = 0 \tag{16}$$

we obtain

$$\left.\begin{aligned} n_i\ddot{X}_i &= -\frac{\mu n_i}{\rho_i^3}X_i + \alpha \\ n_i\ddot{Y}_i &= -\frac{\mu n_i}{\rho_i^3}Y_i + \beta \\ n_i\ddot{Z}_i &= -\frac{\mu n_i}{\rho_i^3}Z_i + \gamma \end{aligned}\right\} \quad i = 1,2,3 \tag{17}$$

From Eq. (17) and

$$\sum_{i=1}^{3}\ddot{X}_i = 0$$

we obtain

$$\alpha\sum_{i=1}^{3}\frac{1}{n_i} = \mu\sum_{i=1}^{3}\frac{X_i}{\rho_i^3}$$

so that

$$\left.\begin{aligned} \alpha &= m_0^2\sum_{1}^{3}\frac{X_i}{\rho_i^3} \\ \beta &= m_0^2\sum_{1}^{3}\frac{Y_i}{\rho_i^3} \\ \gamma &= m_0^2\sum_{1}^{3}\frac{Z_i}{\rho_i^3} \end{aligned}\right\} \tag{18}$$

Equations (17) and (18) yield Eq. (5) when written in vector form.

Reference

1. Deprit, A., and Delie, A., "Systèmes Materiels de Trois Masses Ponctuelles. Mouvement Plan en Coordonnees Symetriques," *Annales de la Societé Scientifique de Bruxelles*, Vol. 75, No. 1, pp. 5–44, 1961.

B. A Critique of the Scalar-Tensor Theory of Relativity, *H. Lass*

Based on remarks of P. M. Dirac concerning the so-called gravitational constant κ, a study was made by Heckmann, Jordan, and Fricke (Ref. 1) in which κ was considered a scalar whose value was dependent on the distribution of matter in Einstein's space–time theory of gravitation.

The same type of theory was rediscovered by Brans and Dicke (Ref. 2). In the Brans–Dicke theory, κ is replaced by $1/\phi$. Moreover, ϕ is not coupled to the Lagrangian of matter, so that the Brans–Dicke theory has the advantage that $T^{ij}_{\ ,j} = 0$, with T^{ij} the energy-momentum tensor, which guarantees that particles move along geodesics. In the former theory, one obtains $(\kappa^2 T^{ij})_{,j} = 0$.

We have found an error in the Brans–Dicke field equations obtained from the variational principle applied to

$$\delta \int \left(\phi R - \frac{\omega}{\phi} g^{\mu\nu} \phi_{,\mu} \phi_{,\nu} + \frac{16\pi}{c^2} L \right) (|g|)^{1/2} d\tau = 0 \qquad (1)$$

with R_{ij} the Ricci tensor, $R = g^{\mu\nu} R_{\mu\nu}$, ω a coupling constant, L the Lagrangian of matter, and $(|g|)^{1/2} d\tau$ the four-dimensional volume element.

Varying ϕ and $\phi_{,\mu}$ yields

$$R - \frac{\omega}{\phi^2} g^{\mu\nu} \phi_{,\mu} \phi_{,\nu} + \frac{2\omega}{\phi} g^{\mu\nu} \phi_{,\mu\nu} = 0 \qquad (2)$$

which agrees with the Brans–Dicke result.

Varying g_{kl} yields the set of equations

$$\left(R_{kl} - \frac{1}{2} g_{kl} R \right) - \frac{1}{\phi} (g_{kl} \Box\phi - \phi_{,kl})$$

$$- \frac{\omega}{\phi^2} \left[\phi_{,k} \phi_{,l} - \frac{1}{2} g_{kl} g^{\mu\nu} \phi_{,\mu} \phi_{,\nu} \right] = - \frac{8\pi}{\phi c^2} T_{kl} \qquad (3)$$

The term

$$- \frac{1}{\phi} (g_{kl} \Box\phi - \phi_{,kl})$$

differs from that of Brans–Dicke in sign only.[1]

If we multiply the terms of Eq. (3) by g^{kl} and add this result to Eq. (2), we obtain

$$\Box\phi = \frac{8\pi T}{c^3 (3 - 2\omega)}, \qquad \omega \neq \frac{3}{2} \qquad (4)$$

with $\Box\phi = g^{\mu\nu} \phi_{,\mu\nu}$, $T = g^{kl} T_{kl}$.

One can then verify that $T^{kl}_{\ ,l} = 0$. Making use of Eqs. (2) and (4), the field Eq. (3) becomes

$$R_{kl} + \frac{1}{\phi} \phi_{,kl} - \frac{\omega}{\phi^2} \phi_{,k} \phi_{,l} = \frac{-8\pi}{c^2 \phi} T_{kl} + g_{kl} \frac{1 - \omega}{3 - 2\omega} \frac{8\pi T}{c^2} \qquad (5)$$

In the absence of matter, $T_{kl} = 0$, $T = 0$, and

$$\left.\begin{aligned} & g^{\mu\nu} \phi_{,\mu\nu} = \Box\phi = 0 \\ & R_{kl} + \frac{1}{\phi} \phi_{,kl} - \frac{\omega}{\phi^2} \phi_{,k} \phi_{,l} = 0 \end{aligned}\right\} \qquad (6)$$

Let us consider now the case of a "uniform" gravitational field due to a uniform distribution of matter in the y–z plane at $x = 0$. Without going into details, it is easily shown that

$$\left.\begin{aligned} ds^2 &= c^2 \exp\left(\frac{2gx}{c^2}\right) dt^2 \\ &\quad - \exp\left(\frac{2g}{c^2} \frac{\omega - 2}{\omega} x\right) dx^2 - dy^2 - dz^2 \\ \phi(x) &= \phi_0 \exp\left(\frac{-2gx}{c^2 \omega}\right) \\ \kappa(x) &= \kappa_0 \exp\left(\frac{2gx}{c^2 \omega}\right) \end{aligned}\right\} \qquad (7)$$

satisfies the field Eq. (6).

[1]The author is indebted to Dr. P. Gottlieb for verifying the conclusions regarding this term.

The case $\omega > 0$ yields a strange result for, as we move away from the source, $\kappa(x)$ increases without bound. If, on the other hand, $\omega < 0$, then $\kappa(x)$ decreases as $x \to \infty$. This is a more satisfactory result from the Machian point of view.

From this elementary example we have come to the conclusion that if $\omega > 0$ then ϕ increases (κ decreases) as one approaches the mass source, while if $\omega < 0$ then ϕ decreases (κ increases) as one approaches the mass source. These observations will be pertinent when we discuss the case of a single mass point (sun).

We take as the line element

$$ds^2 = c^2 e^{\nu(r)}\, dt^2 - e^{\lambda(r)}\, dr^2 - r^2 (d\theta^2 + \sin^2\theta\, d\phi^2) \tag{8}$$

The field Eq. (6) becomes

$$\left.\begin{aligned}
&\frac{1}{2}\nu'' + \frac{1}{4}\nu'^2 - \frac{1}{4}\lambda'\nu' + \frac{\nu'}{r} + \frac{1}{2}\nu'\frac{\phi'}{\phi} = 0\\
&\frac{1}{2}\nu'' + \frac{1}{4}\nu'^2 - \frac{1}{4}\lambda'\nu' - \frac{1}{r}\lambda' =\\
&\qquad -\frac{1}{\phi}\phi'' + \frac{1}{2}\lambda'\frac{\phi'}{\phi} + \frac{\omega}{\phi^2}\phi'^2\\
&1 + \frac{r}{2}(\nu' - \lambda') + \frac{1}{\phi} r\phi' = e^{\lambda}\\
&2r\phi' + r^2\phi'' + \frac{r^2}{2}(\nu' - \lambda')\phi' = 0
\end{aligned}\right\} \tag{9}$$

To obtain the advance in perihelion, we need only determine the unknown constants a, α, β, γ by letting

$$\left.\begin{aligned}
&e^{\lambda} \approx 1 + \frac{a}{c^2 r}, \qquad \lambda = \frac{a}{c^2 r}\\
&\nu = -\frac{2\kappa_0 M}{c^2 r} + \frac{\alpha}{c^4 r^2}\\
&\phi = \phi_0\left(1 + \frac{\beta}{c^2 r} + \frac{\gamma}{c^4 r^2}\right)
\end{aligned}\right\} \tag{10}$$

The value of ϕ_0 is immaterial. The term $-(2\kappa_0 M/c^2 r)$ in $\nu(r)$ is needed to ensure Newtonian motion to a first approximation, with κ_0 the present value of κ due to the presence of matter in the universe.

From the previous equations, one determines that

$$\left.\begin{aligned}
&\alpha = -\kappa_0^2 M^2 + \beta\kappa_0 M - \frac{a}{2}\kappa_0 M\\
&a = 2(\kappa_0 M - \beta)\\
&\gamma = \frac{1}{4}\beta(a + 2\kappa_0 M)\\
&a^2 = 2\beta^2 - 2\alpha - 4\gamma
\end{aligned}\right\} \tag{11}$$

from which it follows that

$$(4 + \omega)\beta^2 - 10\beta\kappa_0 M + 4\kappa_0^2 M^2 = 0 \tag{12}$$

For $\omega \neq -4$, one has

$$\beta = \frac{\kappa_0 M}{4 + \omega}\left[5 \pm (9 - 4\omega)^{1/2}\right] \tag{13}$$

so that $\omega \leq 9/4$.

The advance in the perihelion per revolution can be shown to be

$$\begin{aligned}
\Delta' &= \frac{2\pi\kappa_0 M(a + \kappa_0 M)}{c^2 h^2}\\
&\quad + \frac{\pi}{c^2 h^2}\left[\kappa_0 M(2\kappa_0 M - a) - \alpha - 2\kappa_0^2 M^2\right]\\
&= \frac{6\pi\kappa_0^2 M^2}{c^2 h^2}\left(1 - \frac{2\beta}{3\kappa_0 M}\right)
\end{aligned} \tag{14}$$

by virtue of Eq. (11).

The case $\omega = -\infty$ ($\beta = 0$) yields the Einstein value

$$\Delta = \frac{6\pi\kappa_0^2 M^2}{c^2 h^2}$$

For $0 \leq \omega \leq 9/4$, the minimum value of β is $\beta = (1/2)\kappa_0 M$, and

$$\Delta' = \frac{2}{3}\Delta \tag{15}$$

a significant deviation from Einstein's value. The oblateness of the sun cannot account for this substantial deviation. The same type of conclusion holds for $\omega = -4$, for in this case $\beta = (2/5)\kappa_0 M$ and $\Delta' = (11/15)\Delta$.

Thus $\omega < 0$ if the Brans–Dicke theory is to be consistent with experimental evidence. However, for this case, we have seen that ϕ must increase as r increases (moving away from the source). This requires that $\beta < 0$ in which case $\Delta' > \Delta$. If Dicke's observations regarding the oblateness of the sun is correct, then the final result concerning the advance in the perihelion of Mercury is far in excess of the observational values.

It is conceivable, however, that due to the presence of matter in the universe the value of β is positive even though $\omega < 0$. Now $\omega = -70$ yields $\beta = (2/11)\,\kappa_0 M$, using the negative radical in Eq. (13). Thus $\Delta' = (29/33)\,\Delta$, which means that the oblateness of the sun must essentially contribute about 5″ advance in the perihelion of Mercury per century. It does not seem reasonable that $|\omega|$ is of this magnitude.

References

1. Heckmann, O., Jordan, P., and Fricke, W., Z. *Astrophys.*, Vol. 28, 1951.
2. Brans, C., and Dicke, R. H., "Mach's Principle and a Relativistic Theory of Gravitation," *Phys. Rev.*, Vol. 124, No. 3, 1961.

III. Computation and Analysis

SYSTEMS DIVISION

A. The Design of an Algorithm for Solving Large Linear Least-Squares Systems Using Two Backup Tapes, *R. J. Hanson*

An infinite number of methods exists for solving large linear least-squares problems. This article presents one such method that is easy to follow and should present little difficulty to a programmer in translating it to computer code.

This linear least-squares problem is presented in the form

$$\mathbf{Ax} \cong \mathbf{b} \tag{1}$$

where the $m \times n$ matrix $\mathbf{A}$ is partitioned

$$\mathbf{A} = [\mathbf{A}_1, \cdots, \mathbf{A}_s] \tag{2}$$

Each of the matrices $\mathbf{A}_j$, $j = 1, \cdots, s$, are $n \times k$. Thus, $n = ks$ holds.

The critical hypothesis. The values of n and k are small enough so that there is working space in the machine to hold simultaneously two of the matrices $\mathbf{A}_j$ and the vector $\mathbf{b}$, plus 2 arrays of the same length as $\mathbf{b}$.

The description of the algorithm will be given as a stepwise procedure design for a subroutine.

We assume:

(1) There are two tape units $T(p), p = 1, 2$, rewound and ready for use. The user will specify these units.

(2) The data which constitutes the matrix $\mathbf{A}$ is available by columns on $T(1)$.

(3) The right-hand-side vector is in the machine in a working area called B.

(4) Two working areas W1 and W2 are given. These areas will each hold the data for any pair of $\mathbf{A}_j$, $1 \leq j \leq s$.

What we are doing in this algorithm is constructing an $n \times m$ orthonormal transformation matrix $\mathbf{Q}$ (see SPS 37-50, Vol. III, pp. 65–66) so that

$$\mathbf{QA} = \begin{bmatrix} \overbrace{\mathbf{R}}^{n} \\ \mathbf{0} \end{bmatrix} \begin{matrix} \} n \\ \} m - n \end{matrix} \tag{3}$$

where $\mathbf{R}$ is upper triangular. The matrix $\mathbf{R}$ (for convenience let $\mathbf{R} = [\mathbf{R}_1, \cdots, \mathbf{R}_s]$ where each $\mathbf{R}_i$ is $n \times k$) may be singular or nearly singular (SPS 37-56, Vol. III, pp. 39–41).

We now present the construction of $\mathbf{R}$ in this situation.[1]

Step 1 Set $p := 1, p' := 2, j := 1, \ell := 1$.

2 Bring $\mathbf{A}_\ell$ from $T(p)$ to work area W1.

3 Do forward triangularization on rows j through $j + k - 1$ in W1 (see SPS 37-50, Vol. III).

4 Return to the user so that he can store that portion of $\mathbf{R}$ just available, i.e., $\mathbf{R}_\ell$.

5 Come back from the user here. Write W1 on $T(p')$.

6 For $t = \ell + 1, \cdots, s$, do the following: Bring $\mathbf{A}_t$ from $T(p)$ to W2. Apply those transformations used to make $\mathbf{R}_\ell$ (now in W1) to W2 and to B. Then write area W2 on $T(p')$.

7 If $\ell < s$, set $j := j + k$, exchange the contents of p and p', and set $\ell := \ell + 1$. Rewind $T(p), p = 1, 2$, to the start of $\mathbf{A}_\ell$. Then go to step 2. Else

8 The matrices $\mathbf{R}_j, j = 1, \cdots, s$, are available on units $2, 1, 2, 1, \cdots$ consecutively. $\mathbf{R}_1$ is on the first block of unit 2; $\mathbf{R}_2$ is on the second block of unit 1; $\cdots$ $\mathbf{R}_s$ is on the sth block of unit

$$\left\{1 + \frac{1 - (-1)^s}{2}\right\}$$

For sequential accumulation, go through this procedure again.

You can now do the back substitution to solve for $\mathbf{x}$ if the matrix $\mathbf{A}$ is not nearly singular.

In case the matrix $\mathbf{A}$ is nearly singular, you can store the matrices $\mathbf{R}_\ell$ on a third unit as indicated in step 4. Now go through the same manipulations with your tapes, as in the present algorithm, and reduce $\mathbf{R}$ to row echelon form (see SPS 37-55, Vol. III, pp. 28–30, and SPS 37-56, Vol. III, pp. 39–41). This will record which variables are being solved.

In back substitution you bring the matrices $\mathbf{R}_j$ back into the machine in the order $j = s, \cdots, 1$ to compute the solution to Eq. (1).

[1]The notation $p := b$ means that a symbolic machine area p is to be replaced by the computed value symbolically labelled b.

IV. Space Instruments

SPACE SCIENCES DIVISION

A. Electronics System for a Three-Axis High-Field Vector Helium Magnetometer, *L. L. Lewyn*

1. Introduction

This article describes an electronics system for operation of a three-axis high-field vector helium magnetometer. Design goals for the system included operation to 10 gauss at bandwidths in excess of 10 Hz with full-wave decommutation of the sensor error signal. The system circuit design was implemented utilizing integrated circuit operational amplifiers, logic elements, and analog switches.

2. System Description

A system block diagram is shown in Fig. 1. The 1.083-μ resonant radiation from the helium lamp is circularly polarized and illuminates a lead sulfide detector after passing through the absorption cell. The absorption cell is surrounded by a triaxial set of coils. These coils are driven so that a rotating field Hs of constant amplitude is produced first in the X–Z plane and then in the Y–Z plane. The coil drive waveforms required to produce the sweep field are shown in Fig. 2.

If a small ambient field H is impressed upon the cell, then the lead sulfide detector will yield an output signal given by

$$S = \frac{K}{2}\left[1 + \cos 2wt + \frac{H}{H_s}\cos\phi\,(\cos wt - \cos 3wt)\,\frac{-H}{H_s}\sin\phi\,(\sin wt + \sin 3wt)\right]$$

In this equation, ϕ denotes the angle between the ambient field and the optical axis in the plane of the sweep field, w denotes the frequency of the sweep field which for this instrument is 200 Hz, the terms containing H/H_s represent the error signal caused by the ambient field components in the plane of the sweep field, the cosine term is the error component in the direction of the light axis, and the sine term is the error component normal to the light axis.

The composite error signal is amplified by the preamplifier and transmitted to the signal amplifier. Switches in the signal amplifier then decommutate the error signal. The separate X, Y, and Z error signals are then integrated, filtered, and transmitted to the drive amplifiers. The error signals are combined with the sweep signals in the drive amplifiers. The ac components of the drive signals to the coils produce the constant amplitude rotating sweep field, and the dc components null the field in the absorption cell.

3. Preamplifier and Detector Bias Supply

The preamplifier amplifies the low-level, high-impedance detector signal and removes the sweep signal second harmonic. The preamplifier and sensor bias supply is shown in Fig. 3. The preamplifier was designed by F. Vescelus, JPL. Transistors Q1 and Q2 provide the

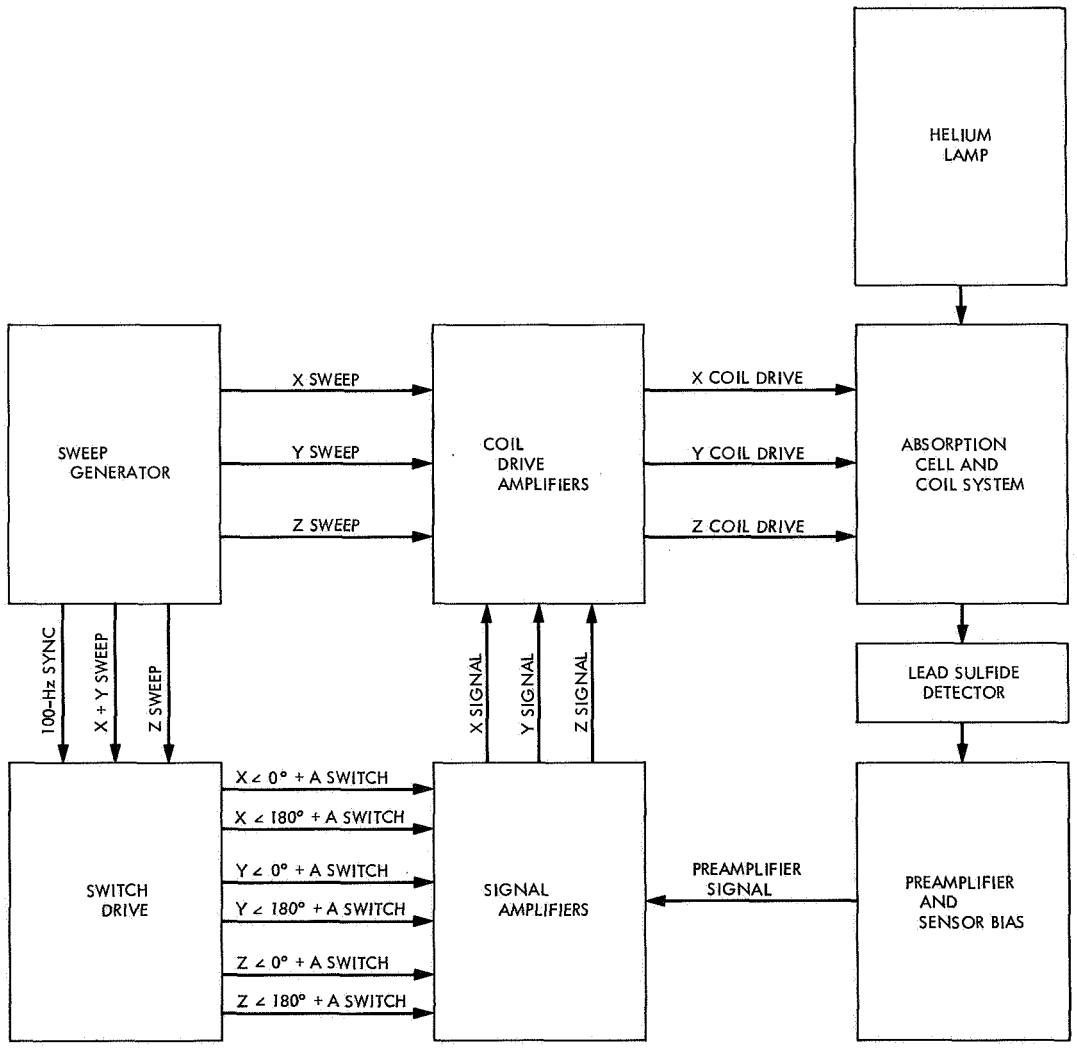


Fig. 1. Magnetometer electronics system

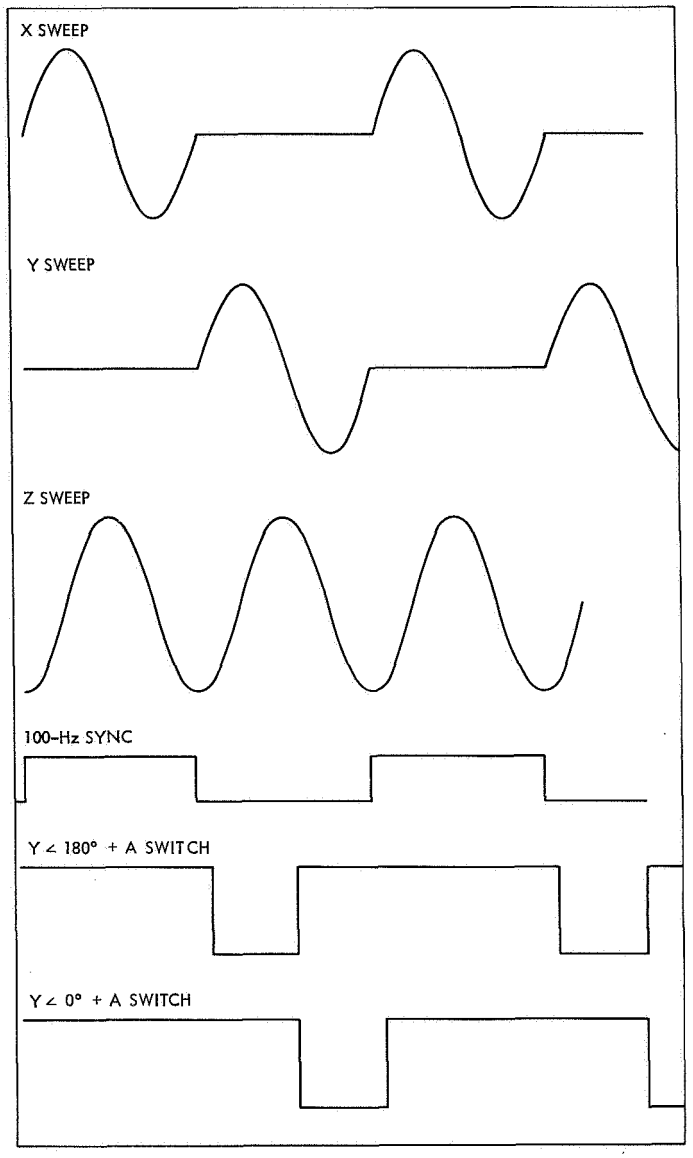


Fig. 2. Drive and switching waveforms

sensor high-impedance ac isolation and small dc voltage drop from the power supplies.

The first stage of the preamplifier is a balanced current-to-voltage amplifier with a forward transfer characteristic v_{out}/i_{in} of 2 MΩ. The balanced first stage provides high common-mode rejection drive to the second stage, which has a voltage gain of 20. The resulting signal is passed through a 400-Hz T-notch filter and the output is buffered by a unity gain isolation amplifier.

4. Signal Amplifiers

The signal amplifiers amplify and perform full-wave decommutation of the preamplifier signals. The decommutated signals are integrated and filtered, then fed to the drive amplifiers for power amplification.

The *Y*-axis signal amplifier is shown in Fig. 4. The first stage of the signal amplifier is a phase splitter with a gain of two. The phase splitter furnishes an in-phase and an inverted signal to drive the switches in the signal chain on each axis.

The *Y*-axis switch drive is shown in Fig. 2. Note that the *X*- and *Y*-axis sweep waveforms appear on alternate sweep cycles. The *Y*-axis switch drive appears only on the sweep cycles when the *Y* axis is being driven. *X* signals are, therefore, prevented from entering the *Y* signal chain. Note also that the *Z* sweep is at quadrature with the *X* and *Y* sweep. Since the *Z* signal is at quadrature with the *Y* signal, it is removed by the synchronous operation of the *Y* switch. The small delay between the start of the *Y* sweep cycle and the operation of the *Y* switch compensates for signal delay due to the accumulation of sensor and amplifier lags.

The switch commutation between in-phase and inverter signal components results in the presentation of a full-wave decommutated signal to the integrating amplifiers. The integration time constant is chosen so that loop gain crossover occurs just above 10 Hz. The output of the integrating amplifiers is passed through two successive notch filters which further attenuate the 100-Hz and 200-Hz switching components in the signal.

5. Coil Drive Amplifiers

A separate drive amplifier is used for each axis of the coil system. Each amplifier performs a weighted summation of the sweep and error signals and provides a current drive to the coil system.

The *Y*-axis drive amplifier is shown in Fig. 5. The drive amplifier is composed of a high-gain integrated operational amplifier input stage followed by a low-gain discrete power stage. The gain and power stages are enclosed by a major feedback loop which provides an operational amplifier summing junction at the input of the integrated circuit. The feedback signal is derived from the 2.5-Ω current sensing resistor in the ground return path of the sweep coil.

6. Sweep Generator

The function of the sweep generator is to provide the drive amplifiers with the required coil sweep waveforms shown in Fig. 2. In addition, the sweep generator provides the switch drive circuits with synchronization signals.

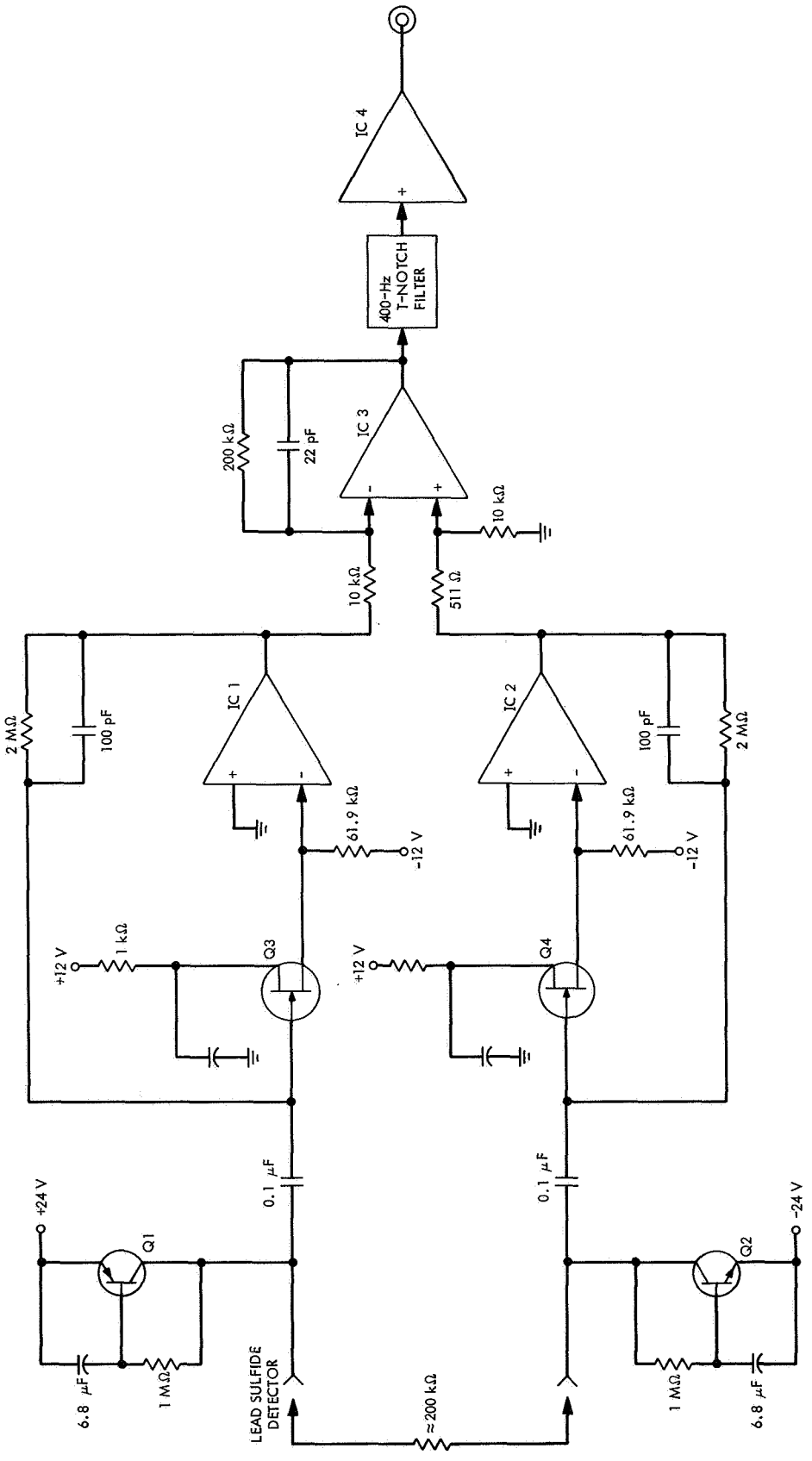


Fig. 3. Preamplifier and sensor bias supply

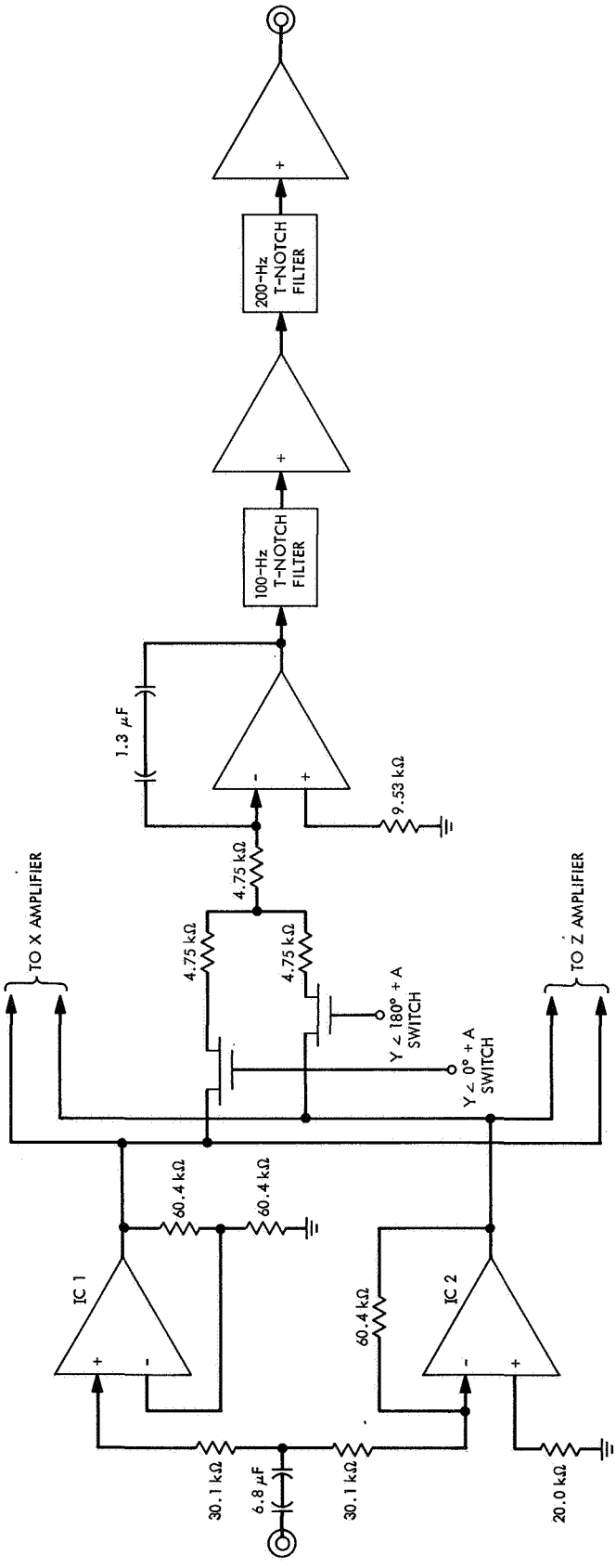


Fig. 4. Y-axis signal amplifier

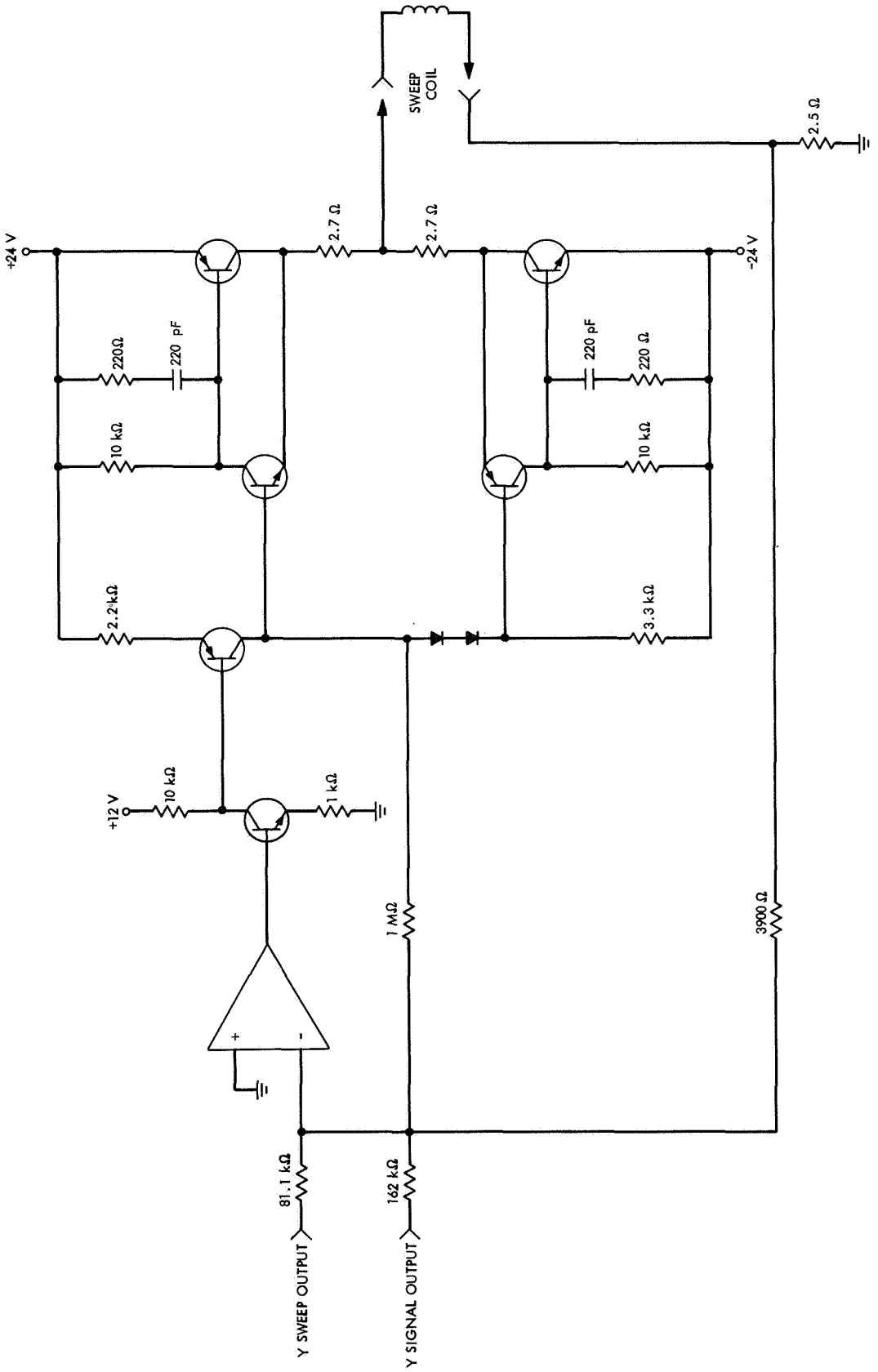


Fig. 5. Y-axis drive amplifier

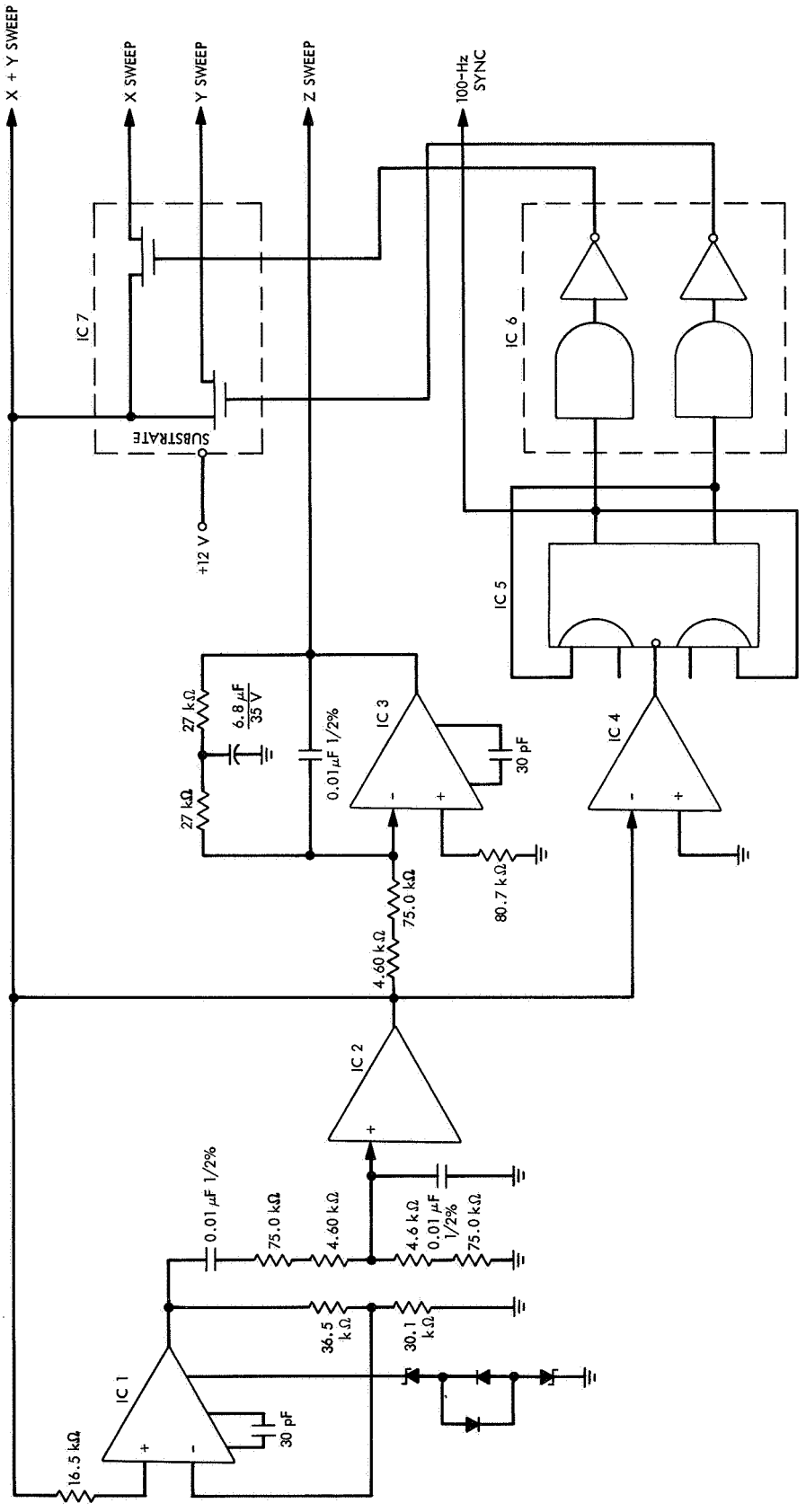


Fig. 6. Sweep generator

The schematic of the sweep generator is shown in Fig. 6. Integrated circuits IC1 and IC2 comprise a Wein bridge oscillator which generates the sinewave required for the X and Y sweep. The sinewave is chopped by the MOS–FET[1] switch (IC7) to provide X and Y sweep on alternate cycles.

The required Z sweep is a sinewave which lags the X and Y sweep by 90 deg. The Z sweep is generated by IC3, which is connected as an integrator with unity gain and 90-deg phase shift at the signal frequency.

A 100-Hz synchronization signal is required to control the decommutation of the error signal in the signal amplifiers. This signal is obtained by detecting the positive zero crossings of the 200-Hz $X + Y$ signal with IC4, which drives the clock input of flip-flop IC5. The flip-flop controls the DTL[2]-to-MOS level shifter IC6 which provides the MOS–FET chopper drive. Consequently, the flip-flop output is always high when the X sweep is passing through the closed switch.

7. Switch Circuitry

The switch circuitry provides the signal amplifier MOS–FET switches with the control required to decommutate the composite error signal from the preamplifier. A typical Y-switch control waveform is shown in Fig. 2. Note that the waveform lags the Y-sweep waveform slightly and is not present during the time that the Y axis is being swept. The switch drive circuitry is shown in Fig. 7.

The $X + Y$ sweep and the Z sweep are summed by the A potentiometer. This phase control provides a sinewave output which has an adjustable phase that may lag the $X + Y$ waveform by as much as 90 deg. The comparator following the phase control has a logic output which changes state as the phase control output crosses zero. The comparator output signals are combined with the 100-Hz flip-flop signal to operate the output DTL-to-MOS level shifters. The flip-flop phasing is controlled by the 100-Hz sync signal from the sweep generator.

The B phase control must provide a signal which lags the Z-sweep waveform from 0 to 90 deg. Therefore, the $X + Y$ signal must be inverted before being summed with the Z sweep. Since the Z sweep is present at all times, no flip-flop control is required.

8. Conclusions

The system electronics has demonstrated stable operation in the laboratory. Frequency response tests indicate that system bandwidths in excess of 10 Hz may be utilized in normal operation. These bandwidths are practical since commonly used narrowband carrier amplifiers have been avoided. Decommutation of the signal at the input of an integrating amplifier provides the system transfer function with a single dominant pole controlling the loop gain crossover. The notch filters which are employed introduce a small decrease in phase margin.

Full-wave sensor error signal decommutation in the system has been achieved using integrated MOS–FET analog switches. Full-wave decommutation provides an inherent signal-to-noise advantage over comparable half-wave chopper–carrier amplifier systems.

The drive electronics has the capability of supplying the power required by the coil system for 10-gauss operation. Summation of the sweep and error signals at the input of the drive amplifier allows the full utilization of its output voltage by the coil system. Further advances in the design of compact high-field coil systems by W. Momsen of JPL will probably eliminate the requirement for a discrete power stage in the drive amplifier.

[1]MOS–FET = metal oxide semiconductor–field-effect transistor.
[2]DTL = diode-transistor logic.

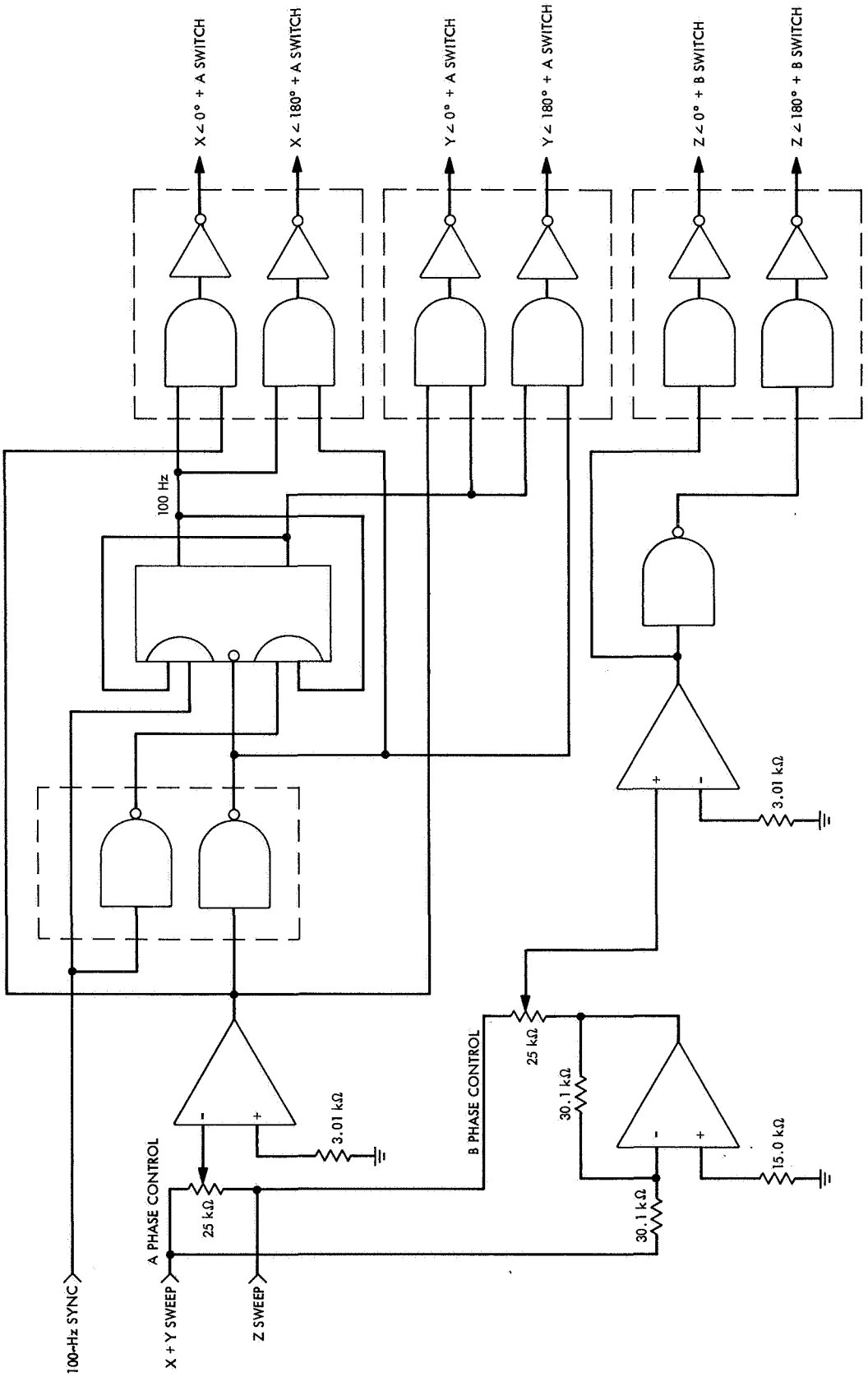


Fig. 7. Switch drive circuitry

V. Lunar and Planetary Sciences

SPACE SCIENCES DIVISION

A. Error Analysis of the Linearization Method in Radiative Transfer Theory, *K. D. Abhyankar*[1] *and A. L. Fymat*

1. Sample Computations

The linearization method of solving the problem of radiative transfer in inhomogeneous atmospheres of arbitrary optical thickness was described in SPS 37-56, Vol. III, pp. 50–52, and an iteration scheme was given in SPS 37-56, Vol. III, pp. 53–55.[2] For illustrating the method, we have made computations of H-functions for isotropically scattering semi-infinite homogeneous atmospheres with albedos for single scattering $\Omega = 0.975$, 0.925, 0.700, and 0.300 from the H-function for a reference atmosphere with $\Omega_0 = 0.900$ (Ref. 1). These cases correspond to the perturbation values: $\omega = +1/12$, $+1/36$, $-2/9$, and $-2/3$, respectively. The results are given in Table 1. The differences between these linear computations and the exact ones are found to be small; hence, the linearization method is quite effective.

[1]Senior Postdoctoral Resident Research Associate of NRC-NASA; on leave of absence from Osmania University, Hyderabad, India.

[2]See Erratum on p. 35 of this volume.

2. Definition of Error

The linearization method consists in expressing the albedo Ω and the X- or H-function in the forms

$$\Omega(\tau') = \Omega_0 [1 + \omega(\tau')], \qquad (\omega << 1)$$

$$V(\tau'; \mu) = V_0(\tau'; \mu)[1 + v(\tau'; \mu)], \qquad (v << 1; V \equiv X, H)$$

V_0 being the V-function for a homogeneous atmosphere of thickness τ' with constant albedo Ω_0. This gives the nonlinear singular integral equation

$$v(\tau; \mu) = \frac{\Omega_0}{V_0(\tau; \mu)} \int_0^1 \left[\int_0^\tau V_0(\tau'; \mu) V_0(\tau'; \mu') \{\omega(\tau') + [1 + \omega(\tau')][v(\tau'; \mu) + v(\tau'; \mu') + v(\tau'; \mu) v(\tau'; \mu')]\} \times \exp\left\{-(\tau - \tau')\left(\frac{1}{\mu} + \frac{1}{\mu'}\right)\right\} d\tau' \right] \Psi(\mu') \frac{d\mu'}{\mu'}$$

We shall denote the solution of this equation by $v_t(\tau; \mu)$ and the corresponding solution of the linearized equation, obtained when neglecting the quadratic term in v,

Table 1. *H*-functions for various homogeneous atmospheres obtained from a reference atmosphere with albedo $\Omega_0 = 0.900$

μ	$\Omega = 0.975, \omega = +1/12$		$\Omega = 0.925, \omega = +1/36$		$\Omega = 0.700, \omega = -2/9$		$\Omega = 0.300, \omega = -2/3$	
	Linear	Exact	Linear	Exact	Linear	Exact	Linear	Exact
0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
0.1	1.2062	1.2111	1.1825	1.1828	1.1112	1.1131	1.0382	1.0399
0.2	1.3587	1.3703	1.3116	1.3123	1.1776	1.1825	1.0562	1.0612
0.3	1.4921	1.5117	1.4213	1.4224	1.2278	1.2364	1.0676	1.0764
0.4	1.6132	1.6414	1.5182	1.5197	1.2674	1.2806	1.0750	1.0881
0.5	1.7234	1.7621	1.6051	1.6073	1.3002	1.3180	1.0800	1.0976
0.6	1.8265	1.8753	1.6843	1.6869	1.3276	1.3501	1.0837	1.1054
0.7	1.9223	1.9822	1.7569	1.7600	1.3508	1.3782	1.0860	1.1120
0.8	2.0123	2.0833	1.8238	1.8274	1.3705	1.4029	1.0878	1.1176
0.9	2.0974	2.1795	1.8860	1.8898	1.3870	1.4250	1.0886	1.1225
1.0	2.1774	2.2710	1.9436	1.9479	1.4026	1.4448	1.0892	1.1268

by $v(\tau;\mu)$. Then, defining the error by $e_v(\tau;\mu) = v_t(\tau;\mu) - v(\tau;\mu)$, we have the equation

$$e_v(\tau;\mu) = \frac{\Omega_0}{V_0(\tau;\mu)} \int_0^1 \left[\int_0^\tau V_0(\tau';\mu) V_0(\tau';\mu') [1 + \omega(\tau')] \times [e_v(\tau';\mu) + e_v(\tau';\mu') + v_t(\tau';\mu) v_t(\tau';\mu')] \times \exp\left\{ -(\tau - \tau')\left(\frac{1}{\mu} + \frac{1}{\mu'}\right)\right\} d\tau' \right] \Psi(\mu') \frac{d\mu'}{\mu'}$$

for determining the error caused by the process of linearization. We shall also be interested in the relative errors $\{e_v(\tau;\mu)/v(\tau;\mu)\}$ and $\{e_V(\tau;\mu)/V(\tau;\mu)\}$, where $e_V = V_{\text{exact}} - V_{\text{linear}}$.

3. Evaluation of the Error for a Semi-infinite Atmosphere

Let $s_1 \leqq \omega(\tau') \leqq s_2$, where both the constants s_1 and s_2 can lie in the range -1 to $(1 - \Omega_0)/\Omega_0$. Now, for fixed μ, the absolute value of the error $e_h(\tau;\mu)$ will be maximum at some value of τ, $\tau = \tau_0$, say. We then expect that we can find a semi-infinite homogeneous atmosphere with $\omega(\tau') \equiv s$, where s is some constant between s_1 and s_2 not necessarily equal to $\omega(\tau_0)$, such that it will reproduce a constant error $e_h(\mu) = e_h(\tau_0;\mu)$ throughout the atmosphere. For such an atmosphere, since $h_t(\tau;\mu) = h_t(\mu)$ and $h(\tau;\mu) = h(\mu)$ are also independent of τ, the error equation becomes

$$e_h(\mu) = (1 + s)\Omega_0 \int_0^1 H_0(\mu')\Psi(\mu') \times [e_h(\mu) + e_h(\mu') + h_t(\mu) h_t(\mu')] \frac{\mu\, d\mu'}{\mu + \mu'}$$

Numerical computations indicate that $e_h(\mu)$ is always positive and varies roughly linearly from zero to $e_h(1)$ as μ increases from zero to unity. Hence, taking the mean value of $e_h(\mu')$ to be $\frac{1}{2} e_h(1)$, we get[3]

$$\left\{\frac{e_h(\mu)}{h_t(\mu)}\right\}_{\max} = \frac{e_h(1)}{h_t(1)} = \frac{(1 + s)\Omega_0 j_t(1)}{1 - \frac{3}{2}(1 + s) f_1(1)}$$

where

$$f_1(\mu) = \Omega_0 \int_0^1 H_0(\mu')\Psi(\mu') \frac{\mu\, d\mu'}{\mu + \mu'} = 1 - H_0^{-1}(\mu)$$

and

$$j_t(\mu) = \int_0^1 H_0(\mu') h_t(\mu')\Psi(\mu') \frac{\mu\, d\mu'}{\mu + \mu'}$$

$$= \frac{1 - H_t^{-1}(\mu)}{\Omega_0(1 + s)} - \frac{1 - H_0^{-1}(\mu)}{\Omega_0}$$

The maximum relative error in H is thus given by

$$\left\{\frac{e_H(\mu)}{H_t(\mu)}\right\}_{\max} = \frac{e_H(1)}{H_t(1)} = \frac{e_h(1)}{h_t(1)}\left[\frac{H_t(1) - H_0(1)}{H_t(1)}\right]$$

A similar analysis for a finite atmosphere seems to be more difficult.

[3]Fymat, A. L., and Abhyankar, K. D., Paper II, *Astrophys. J.*, Oct. 1969 (in press).

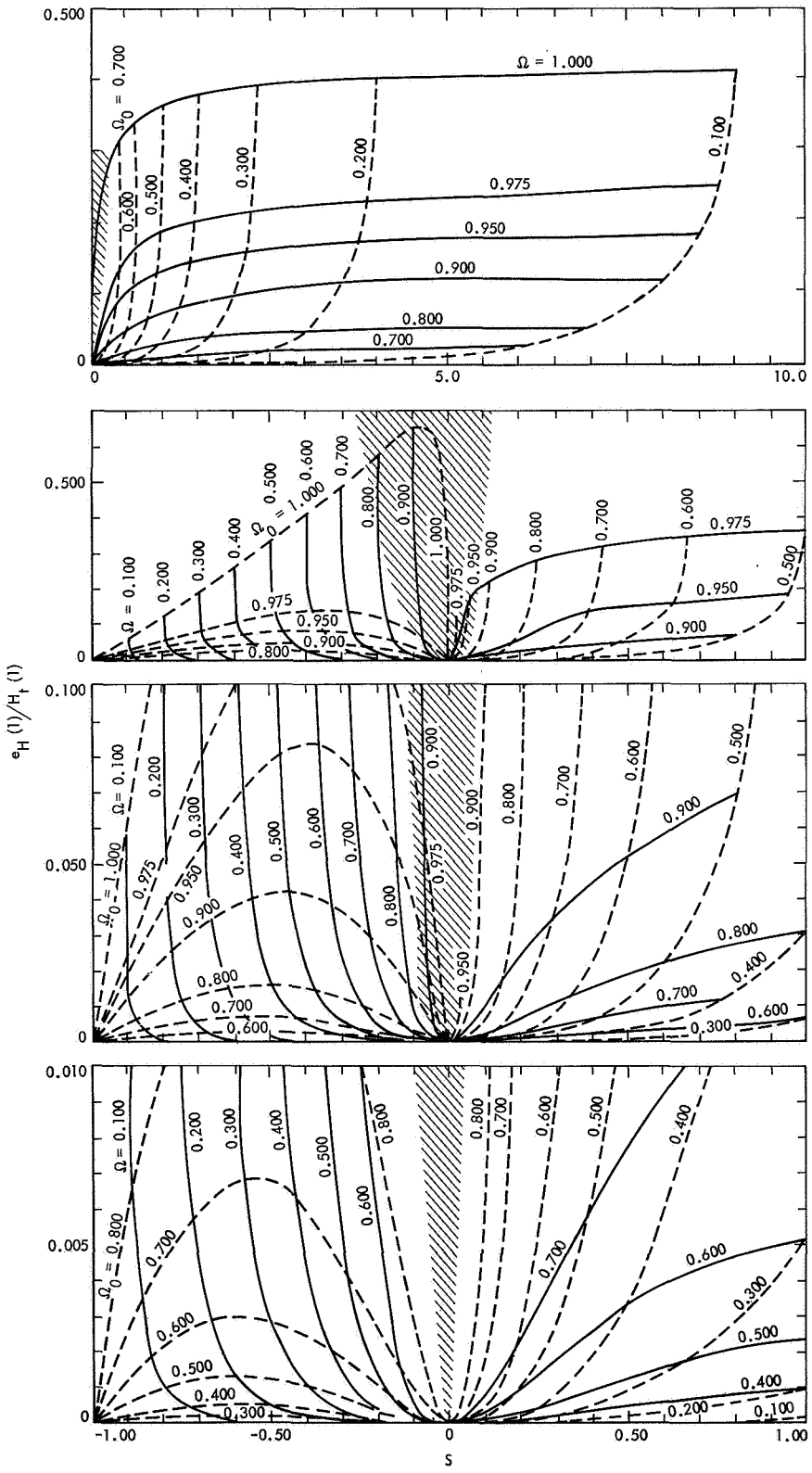


Fig. 1. Maximum relative error in $H(\mu)$ as a function of s for various values of Ω_0

4. Error Diagrams for Isotropic Scattering

The H-functions for homogeneous atmospheres which exhibit isotropic scattering, $\Psi(\mu) = \frac{1}{2}$, have been tabulated for several values of the albedo for single scattering (Ref. 1). Using these tables, we have computed the quantity $[e_H(1)/H_t(1)]$ for various combinations of Ω_0 and s. The results are shown in Fig. 1; hatched areas indicate regions where the linear solution is not convergent. It can be seen that the maximum errors found in the sample cases correspond with those indicated by this figure. It also shows that, in certain cases, the linear equations give fairly accurate results even when the perturbation is large. For example, from $\Omega_0 = 0.2$, we can get the solution for $\Omega = 0.4$ accurate to within 0.1%, even though the so-called "perturbation," s, takes the large value of 1.0.

The diagrams of Fig. 1 will be useful for estimating the maximum relative error in $H(\tau; \mu)$ to be expected when applying the linearization method to any particular problem. For example, consider a semi-infinite inhomogeneous atmosphere with $0.1 \leqq \Omega(\tau) \leqq 0.8$; if Ω_0 is taken equal to 0.4, 0.6, or 0.8, say, the corresponding maximum errors in $H(\mu)$ would be 3.1, 0.3, and 1.6%, respectively. Hence, the best choice of Ω_0 would be 0.6. In general, the errors are smaller when the perturbations are negative so that it is always advantageous to select Ω_0 accordingly.

In the case of finite atmospheres, the error diagrams must be constructed empirically from numerical computations.

Reference

1. Chandrasekhar, S., *Radiative Transfer*, p. 125. Clarendon Press, Oxford, London, 1950.

B. Effect of Linearization of the Transfer Problem on the Computations of the Emergent Radiation Field, *A. L. Fymat and K. D. Abhyankar*[4]

1. Parameters of the Radiation Field

A planetary atmosphere, usually considered to be plane-parallel stratified, is illuminated from above by a monodirectional beam of solar radiation of flux

$$\pi F \delta(\mu - \mu_0)\, \delta(\phi - \phi_0)$$

μ being the cosine of the zenith angle and ϕ the azimuth angle. The reflected and transmitted radiations at the atmospheric boundaries are then given by

$$I_{\text{ref}}(\mu, \phi) = \frac{F}{4\mu} \sum_m S^{(m)}(\mu, \mu_0) \cos[m(\phi - \phi_0)],$$

$$(m = 0, 1, 2, \cdots)$$

and

$$I_{\text{trans}}(\mu, \phi) = \frac{F}{4\mu} \sum_m T^{(m)}(\mu, \mu_0) \cos[m(\phi - \phi_0)]$$

where the scattering functions S and T (for all values of m) depend on the optical thickness of the atmospheric slab. They are known as the reflection and transmission functions of the atmosphere. In the case of an inhomogeneous atmosphere in which the albedo for single scattering $\Omega(\tau)$ varies with optical depth, additional scattering functions S^* and T^* should be introduced when the illumination is from below. The solutions of the equations for all these scattering functions are simplified by means of the auxiliary functions X, Y, X^*, Y^* for finite atmospheres and an H-function for a semi-infinite atmosphere (Ref. 1). We have the relations

$$S(\tau; \mu, \mu_0) = \int_0^\tau X(\tau'; \mu) X(\tau'; \mu_0) \times \exp\left\{-(\tau - \tau')\left(\frac{1}{\mu} + \frac{1}{\mu_0}\right)\right\} \Omega(\tau')\, d\tau'$$

$$T(\tau; \mu, \mu_0) = \int_0^\tau Y(\tau'; \mu) X(\tau'; \mu_0) \times \exp\{-(\tau - \tau')/\mu_0\} \Omega(\tau')\, d\tau'$$

and two similar equations for S^* and T^* which are obtained by substituting X^* and Y^* in place of X and Y, respectively. In the case of a semi-infinite atmosphere, Y vanishes and X becomes the H-function.

In this article, we shall investigate the errors in the computations of the scattering functions and of the radiation field which result from our linearization process (SPS 37-56, Vol. III, pp. 50–52).

2. Consideration of the Scattering Functions

In *Section A* of this chapter, the results of computations of the H-functions were given for isotropically scattering homogeneous semi-infinite atmospheres with albedos $\Omega = 0.975$, 0.925, 0.700, 0.300 obtained from those for a reference atmosphere with albedo $\Omega_0 = 0.900$. These

[4]Senior Postdoctoral Resident Research Associate of NRC-NASA; on leave of absence from Osmania University, Hyderabad, India.

values were used to compute the reflection functions shown in Table 2 by means of the following formula appropriate for a homogeneous atmosphere (Ref. 1):

$$\left(\frac{1}{\mu}+\frac{1}{\mu_0}\right) S(\mu,\mu_0) = \Omega H(\mu) H(\mu_0)$$

Table 2. Reflection functions $S(\mu, \mu_0)$*

μ	$\mu_0 = 0.2$	$\mu_0 = 0.4$	$\mu_0 = 0.6$	$\mu_0 = 0.8$	$\mu_0 = 1.0$
			$\Omega = 0.300$		
0.2	0.03347 0.03378	0.04541	0.05151	0.05515	0.05752
0.4	0.04619	0.06934 0.07104	0.08388	0.09355	0.10036
0.6	0.05279	0.08660	0.10570 0.10997	0.12125	0.13280
0.8	0.05693	0.09729	0.12707	0.14200 0.14988	0.15797
1.0	0.05979	0.10509	0.14013	0.16790	0.17796 0.19046
			$\Omega = 0.700$		
0.2	0.09707 0.09788	0.13930	0.16416	0.18076	0.19270
0.4	0.14133	0.22488 0.22959	0.28268	0.32424	0.35553
0.6	0.16763	0.29046	0.37013 0.38279	0.43668	0.48880
0.8	0.18580	0.33537	0.45459	0.52592 0.55107	0.59804
1.0	0.19933	0.37003	0.51203	0.63058	0.68856 0.73059
			$\Omega = 0.925$		
0.2	0.15913 0.15929	0.24559	0.30651	0.35403	0.39301
0.4	0.24596	0.42641 0.42726	0.56768	0.68300	0.77984
0.6	0.30715	0.56912	0.78724 0.78965	0.97421	1.13553
0.8	0.35492	0.68503	0.97763	1.23069 1.23558	1.45725
1.0	0.39409	0.78233	1.13980	1.46338	1.74714 1.75486
			$\Omega = 0.975$		
0.2	0.17999 0.18308	0.28494	0.36295	0.42652	0.48075
0.4	0.29239	0.50747 0.52537	0.68948	0.84402	0.97850
0.6	0.37582	0.72028	0.97581 1.02866	1.22868	1.45409
0.8	0.44533	0.88908	1.30600	1.57927 1.69264	1.89867
1.0	0.50571	1.03839	1.55712	2.05017	2.31129 2.51423

*Computed values on right of diagonal divider, true values on left.

Because of the symmetry of these functions with regard to μ and μ_0, the computed values are entered in the table on the right side of the diagonal while the true values are given on the left.

The relative errors in $S(\mu, \mu_0)$ are shown in Fig. 2. They are usually found to be much smaller than the maximum perturbation; for example, $\{\Delta S/S\}_{max} = 0.066$, 0.058, and 0.0044 for $\omega_{max} = -0.667$, -0.222, and 0.0278, respectively. Only for the large positive perturbation $\omega_{max} = 0.0833$ does the relative error have a comparable value of 0.0807. In general, the effect of linearization is to underestimate the reflection function of the atmosphere. The discrepancy is larger for increasing values of the direction cosines of the incident and reflected beams.

3. Consideration of the Emergent Radiation Field

With the tabulated values of the reflection functions, the reflected radiation was computed from the relation of definition. The errors in the field intensities in units

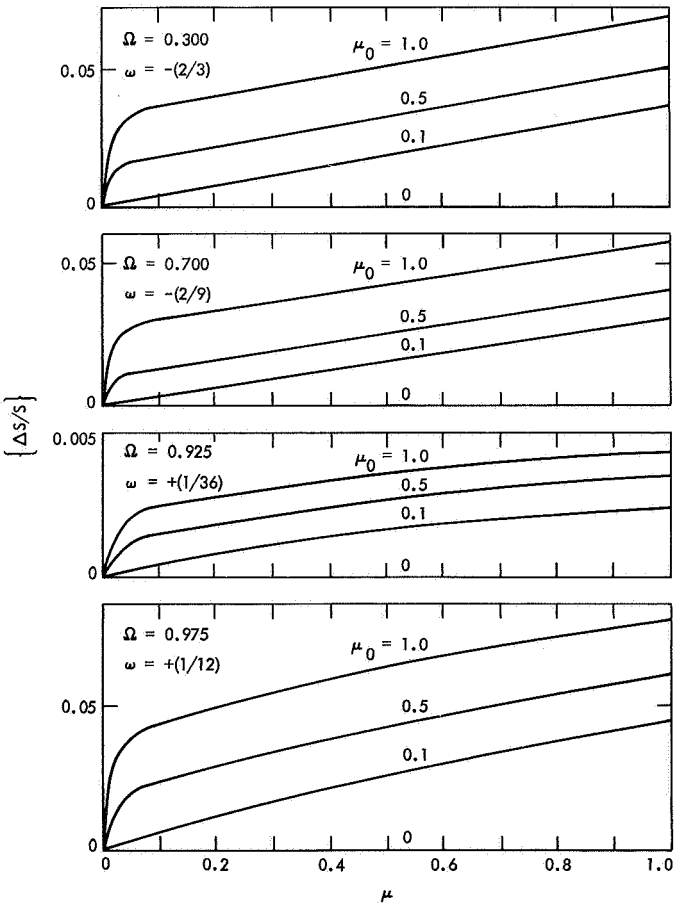


Fig. 2. Relative errors in reflection functions

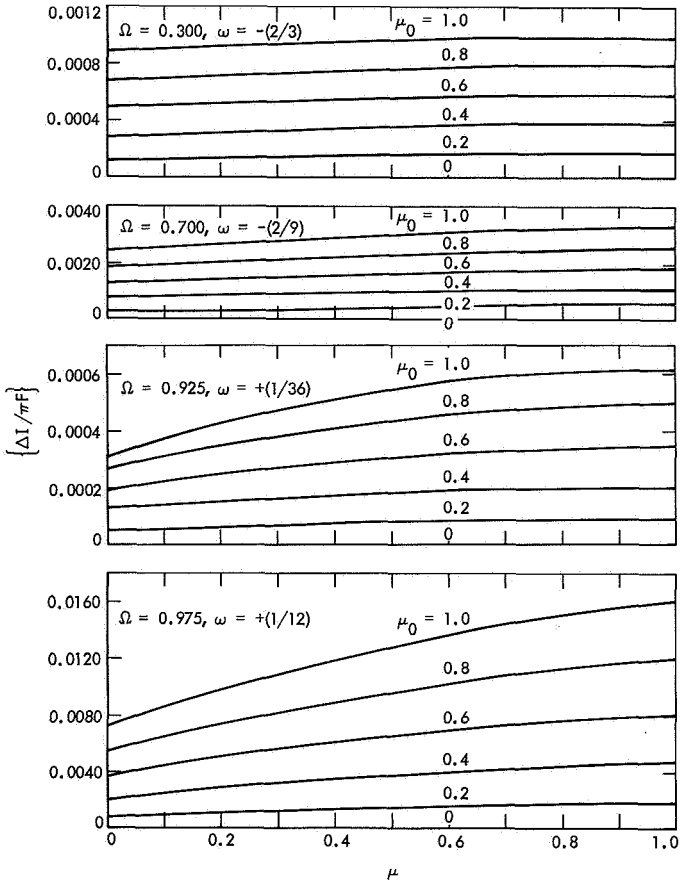


Fig. 3. Errors in field intensities

of the incident flux πF are shown in Fig. 3. These curves are similar to those for S in their variations with μ and μ_0, but the absolute errors involved are always smaller. In the cases considered, their maximum values are, respectively, 0.10, 0.34, 0.06, and 1.6% in units of πF. Thus, the effects of linearization on intensity are generally small.

Now, in the theory of line profiles in planetary spectra, it is the reflection function which is relevant. If $r(\nu)$ is the line depth defined by

$$r(\nu) = \frac{S_{\text{continuum}} - S_{\text{line}}}{S_{\text{continuum}}} = \frac{\delta S}{S}$$

we can write for the relative error in r:

$$\frac{\Delta r}{r} \simeq -\frac{\Delta S}{S}$$

Since the relative error in S caused by the linearization process is small, as shown by the computations reported here, the line profiles obtained in this way will closely fit the true profiles for the inhomogeneous atmosphere under investigation. In some other problems, like the calculation of a planet's phase curve, it is the intensity that must be considered. As seen above, the linearization method gives fairly accurate results in this case also.

Reference

1. Chandrasekhar, S., *Radiative Transfer*. Clarendon Press, Oxford, London, 1950.

C. Lunar Mascons: Refined Calculation of the Model Gravity Profile for Mare Serenitatis,

J. E. Conel, G. B. Holstrom, and E. Abbott

In Ref. 1 Conel and Holstrom presented gravity profiles for a near-surface disk-model of the Mare Serenitatis mass concentration (mascon). Good agreement was obtained between that model and the normalized gravity data of Muller and Sjogren (Ref. 2). These calculations were approximate in the sense that: (1) curvature of the moon's surface was neglected, (2) the gravity effects of topographically high areas ringing the mare were not accounted for, (3) the line-of-sight doppler values were assumed to represent the local vertical component of gravity (this assumption becomes especially important for the northern portions of the profile), and (4) the profile was matched to so-called normalized gravity values, i.e., those values obtained after a "free-air" correction of the raw accelerations to a surface at an altitude of 100 km.

By modifying the original gravity program (Ref. 3) to calculate three orthogonal components of acceleration, and by constructing a more elaborate model including topography and curvature, we have been able to account for effects arising from all of these assumptions in our model calculations. The result is that fit of the plate model to the observational data is sensibly unaffected by such refinements. Comparison with a deep spherical mass distribution also remains favorable. In fact, it appears that the plate and sphere models will remain gravitationally indistinguishable regardless of the altitude of observation unless (1) accuracy of the observational data is improved, or (2) the plate model has a distinctive geometry.

To account for curvature, the anomalous mass distribution is referred to a rectangular coordinate system with origin located at the surface at the center of the model and at latitude θ_0, longitude ϕ_0. The (x, y, z) coordinates of a model point located at (R, θ_1, ϕ_1) are then given by

$$x = \pm R\,[\{\sin[\cos^{-1}(\cos\theta_0\cos\theta_1\cos(\phi_0-\phi_1) + \sin\theta_0\sin\theta_1)]\}^2 - \sin^2(\theta_0-\theta_1)]^{1/2} \quad (1a)$$

$$y = R\sin(\theta_1-\theta_0) \quad (1b)$$

$$z = H(\theta_1,\phi_1)\,[\cos\theta_0\cos\theta_1\cos(\phi_0-\phi_1) + \sin\theta_0\sin\theta_1] - R\,\{1 - [\cos\theta_0\cos\theta_1\cos(\phi_0-\phi_1) + \sin\theta_0\sin\theta_1]\} \quad (1c)$$

where $R = 1738$ km, the radius of an assumed spherical lunar reference surface, and $H(\theta_1, \phi_1)$ is spacecraft altitude above this surface at (θ_1, ϕ_1). The axes are chosen such that

$$+x \equiv \text{east}$$
$$+y \equiv \text{north}$$
$$+z \equiv \text{up}$$

In Eq. (1a)

$$x > 0 \text{ if } \phi_1 > \phi_0$$

and

$$x < 0 \text{ if } \phi_1 < \phi_0$$

The transformation formula relating that component of acceleration in the direction of earth, g_E, to the components X, Y, and Z in the (x, y, z) system arising from an anomalous mass distribution is

$$g_E = -X\sin\phi + Y\sin\theta\cos\phi - Z\cos\theta\cos\phi \quad (2)$$

The sign conventions on the vectorial quantities in Eq. (2) are

x-component = positive if directed eastward

y-component = positive if directed northward

z-component = positive if directed upward

The model and resulting north–south gravity profiles across Mare Serenitatis are shown in Fig. 4. Average elevations of the rim deposits for the southern leg of our profile was constructed from the Aeronautical Chart and Information Center (ACIC) charts. There are very few relative elevations given for the northern or eastern sectors bordering the mare. Consequently, to simplify the computations, the model was given axial symmetry about the approximate geometric center of the mare (28°N lat, 18°E lon). Rim deposits surrounding the basin actually do not show such symmetry and are discontinuous or absent in places; the model thus adequately accounts for gravitational effects of the topography. Profiles in Fig. 4 are given for a rim deposit density of 2.5 g/cm³. The density contrast between fill and rim deposits was taken to be 1.0 g/cm³, with the maximum anomaly 250 mgals.

From examination of the measured and computed profiles we see that inclusion of topography, curvature, spacecraft elevation, as well as accounting for the actual component of acceleration along the line of sight, has not materially affected the fit of our model curves to the actual data. Comparison with data for the deep sphere also remains about the same.

From Fig. 4, it is also seen that the position of the maximum anomaly is offset approximately 80 km south of the center of the model. This is a geometrical effect arising from two causes. First, both horizontal and vertical components of gravity from the anomalous mass are involved in the line-of-sight acceleration data. For a nearly symmetric model, a profile of the vertical component through the center is essentially symmetrical about the center of gravity. On the other hand, the horizontal component is asymmetrical; it changes sign at the center and by symmetry is nearly zero at the origin. (The situation can be understood in a simple way by considering the case of a spherical mass distribution.) Since both components are combined in the actual data, the maximum in the profile is displaced from the center of mass in a direction corresponding to an *acceleration* of the spacecraft, i.e., to the south, for a south-to-north orbital track. It is easily seen from such arguments that the displacement is zero for a mass at the equator and maximum for a mass at the limbs or poles. In the latter instance, profiles for symmetrical objects can be expected to show pronounced asymmetry, with the acceleration changing sign over the center of mass. Second, the spacecraft height varies along the orbital track. Thus, the closest approach to the center of mass is not necessarily vertically above it.

Since the geometric forms of the surface plate and deep sphere models are so dissimilar, it might be expected that near-surface gravity observations would provide a distinction between the two. In Fig. 5 we present profiles of the vertical component gravity for the plate model ($\Delta\rho = 1.0$ g/cm³) and sphere at depth 200 km for equivalent maximum anomaly. Topography and curvature are neglected. Except for detailed structures in the plate curve, which can be definitely attributed to abrupt changes in thickness of the model, the curves are quite similar. We could eliminate these structures to any desired extent by increasing the number of plates indefinitely, and thus produce profiles quite similar to those

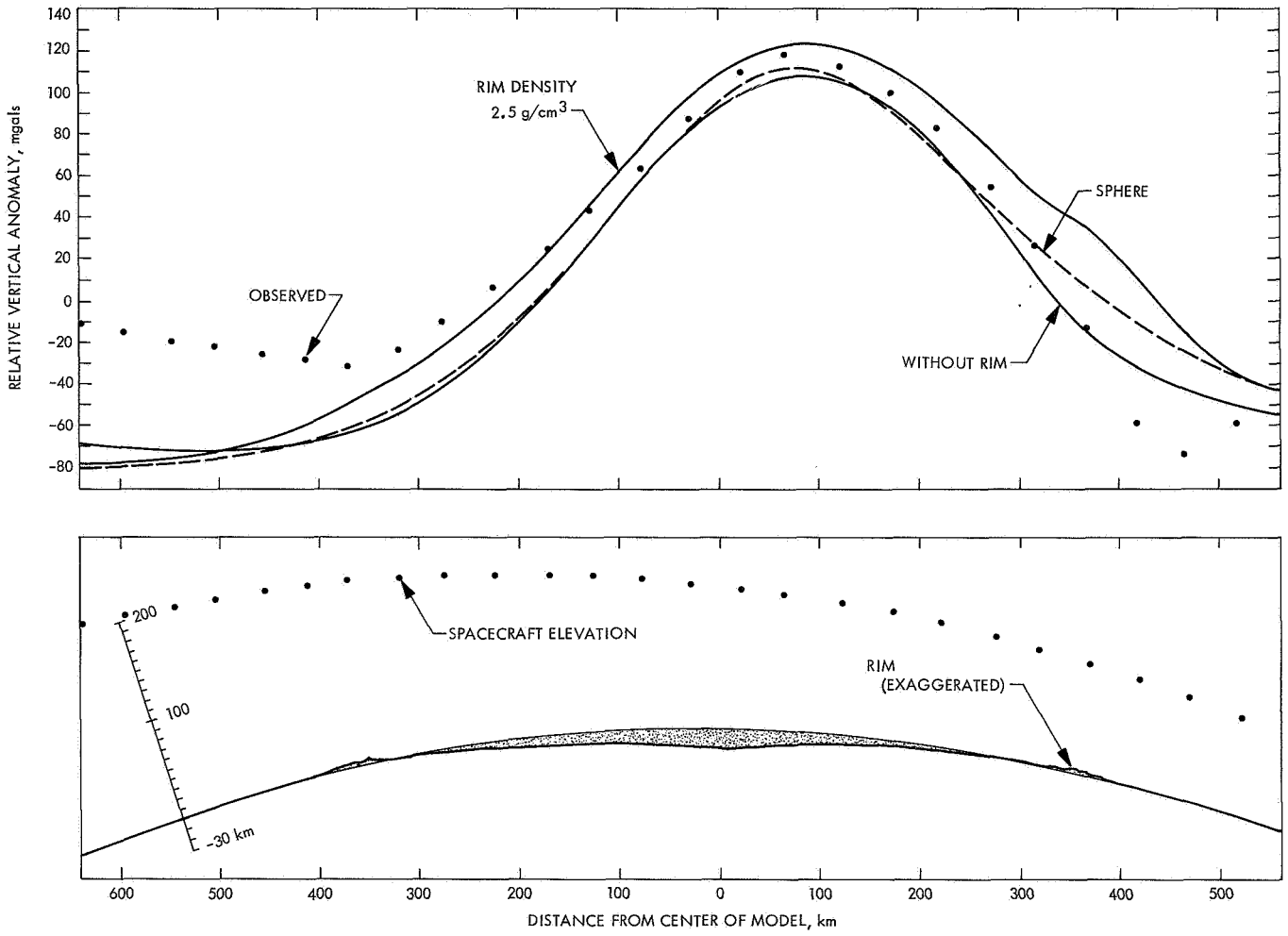


Fig. 4. Comparison of observed and computed profiles for the curved disk model

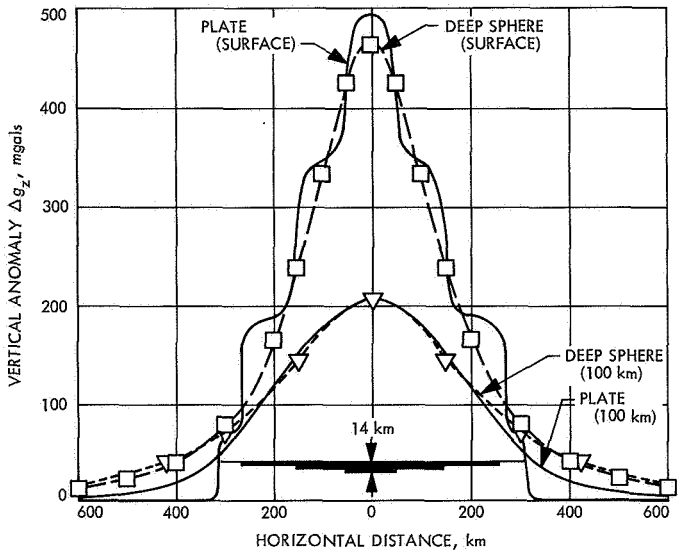


Fig. 5. Theoretical gravity profiles for plate and deep sphere models for spacecraft altitudes of 0 and 100 km

for deeply buried masses. These two models can be expected to show similar gravity effects at the surface, with the maximum anomaly from each differing by about 30 mgals. Thus, precise observations at both elevations would be required to distinguish between the surface-plate and deep-sphere models on this basis. With present accuracy of the data, they would remain gravitationally indistinguishable, unless there are characteristic features in the profiles of the type indicated.

References

1. Conel, J. E., and Holstrom, G. B., "Lunar Mascons: A Near-Surface Interpretation," *Science*, Vol. 162, pp. 1403–1405, 1968.
2. Muller, P. M., and Sjogren, W. L., "Mascons: Lunar Mass Concentrations," *Science*, Vol. 161, pp. 680–684, 1968.
3. Talwani, M., and Ewing, M., "Rapid Computation of Gravitational Attraction of Three-Dimensional Bodies of Arbitrary Shape," *Geophysics*, Vol. 25, pp. 203–225, 1960.

D. Retention and Release of Ion-Implanted Argon,

H. C. Lord, III

An investigation was conducted to determine the effect of substrate composition and pre-irradiation history on the thermal release profiles and retention coefficients of argon injected by ion bombardment into silicates and iron. The samples consisted of polished 1-cm square slabs of the minerals forsterite, enstatite, oligoclase, the glass obsidian, and cold-rolled steel (<1% C). All samples were irradiated with 40-keV Ar^+ ions, but some were pre-irradiated with either H^+, He^+, or both (see Table 3). The procedure for sample preparation and analysis is described in Ref. 1.

The results of the gas analyses are shown in Table 4. The concentrations of evolved Ar/cm² listed in the table

Table 3. Incident ion doses and maximum current densities[a]

Sample	1-keV H^+	4-keV He^+	40-keV Ar^+
A			2.9×10^{17} cm^{-2} (70 μA/cm²)
B			2.8×10^{16} cm^{-2} (30 μA/cm²)
C	1.2×10^{17} cm^{-2} (2.3 μA/cm²)		2.7×10^{16} cm^{-2} (40 μA/cm²)
D		1.5×10^{17} cm^{-2} (1.6 μA/cm²)	3.1×10^{16} cm^{-2} (50 μA/cm²)
E	1.1×10^{17} cm^{-2} (1.6 μA/cm²)	1.5×10^{16} cm^{-2} (1.3 μA/cm²)	3.1×10^{16} cm^{-2} (50 μA/cm²)

[a]Maximum current density values in parentheses.

Table 4. Results of gas analyses

Sample	Incident Ar^+/cm²	Evolved Ar/cm²	Retention coefficient	
			Apparent	Actual[a]
Oligoclase				
A	2.9×10^{17}	4.8×10^{16}	0.17	0.09
B	2.8×10^{16}	2.5×10^{16}	0.89	0.47
C	2.7×10^{16}	2.6×10^{16}	0.95	0.51
D	3.1×10^{16}	2.4×10^{16}	0.77	0.41
E	3.1×10^{16}	2.6×10^{16}	0.82	0.44
Obsidian				
A	2.9×10^{17}	3.2×10^{16}	0.11	0.059
B	2.8×10^{16}	1.3×10^{16}	0.45	0.24
C	2.7×10^{16}	2.2×10^{16}	0.82	0.44
Enstatite				
A	2.9×10^{17}	3.1×10^{16}	0.11	0.059
B	2.8×10^{16}	$>1.3 \times 10^{16}$	>0.48	>0.26
C	2.7×10^{16}	$>2.1 \times 10^{16}$	>0.78	>0.42
Forsterite				
A	2.9×10^{17}	4.1×10^{16}	0.15	0.080
B	2.8×10^{16}	3.2×10^{16}	1.12	0.60
C	2.7×10^{16}	2.1×10^{16}	0.77	0.41
Iron				
A	2.9×10^{17}	2.2×10^{16}	0.078	0.042
B	2.8×10^{16}	1.3×10^{16}	0.45	0.24
C	2.7×10^{16}	2.4×10^{16}	0.86	0.46

[a]Calculated from apparent retention coefficient, taking into account the neutralization charge exchange in the ion beam.

are the sums of the individual concentration measurements at each temperature step. The apparent retention coefficient is merely the measured ratio evolved/incident Ar/cm^2.

The temperature release profile of argon from oligoclase for the five different irradiations is shown in Fig. 6. All five samples show a primary low-temperature maximum in the differential gas release curve. Irradiations C, D, and E (Table 3), which were irradiated with more than one ion, show a secondary higher temperature maximum. Irradiation A, which is at the highest integrated ion flux, has its low-temperature maximum at a higher temperature (400–600° C) than the lower integrated flux irradiations (200–400° C). This suggests increased importance, with irradiation time, of the penetrating tail of the range distribution. Davies and Jespersgård (Ref. 2) have suggested that injected atoms ending up in interstitial sites diffuse rapidly through the lattice until they are trapped by a defect. As has been shown above, surface sputtering will erode the sample, continuously exposing a new surface layer. An equilibrium gas concentration is soon established near the surface, but the gas concentration in the penetrating tail continues to grow.

The effect of substrate composition on argon release is displayed in Fig. 7. These gas release results are for the higher dose (2.9×10^{17} Ar^+/cm^2) irradiation. Obsidian, a partly recrystallized volcanic glass, released the retained gas at very low temperatures (100–300° C). However, the calculated retention coefficient is identical to that of the enstatite sample (which shows a very high temperature release maximum) and greater than that for the iron sample, indicating that there could not have been substantially greater amounts of room-temperature argon loss from the obsidian.

A bi-modal distribution for the argon release is seen in all the samples except the oligoclase. For the iron sample, the two maxima are approximately the same

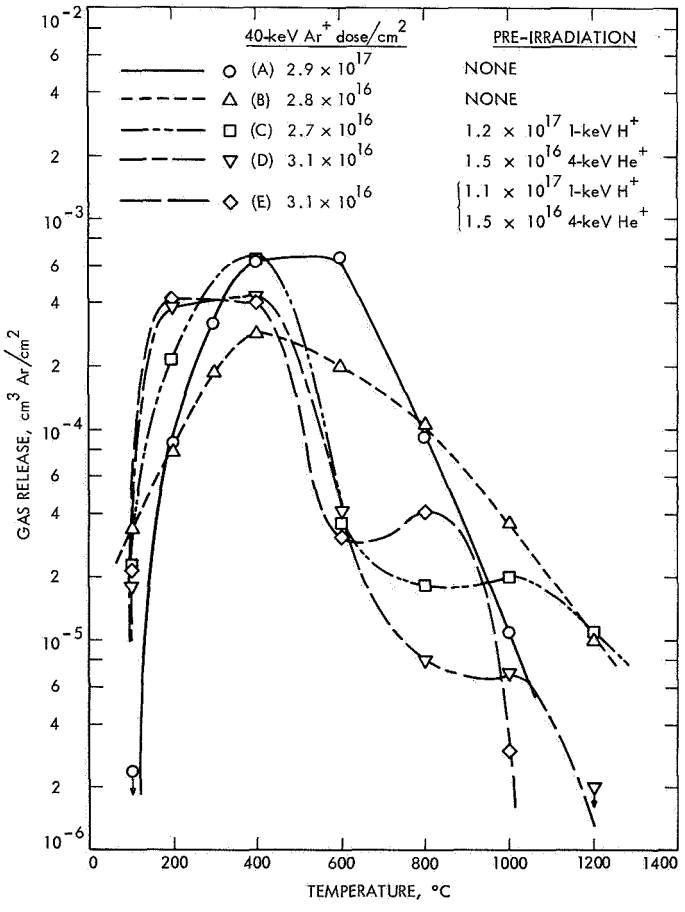


Fig. 6. Differential gas release curves for 40-keV argon ions injected into oligoclase at different irradiations

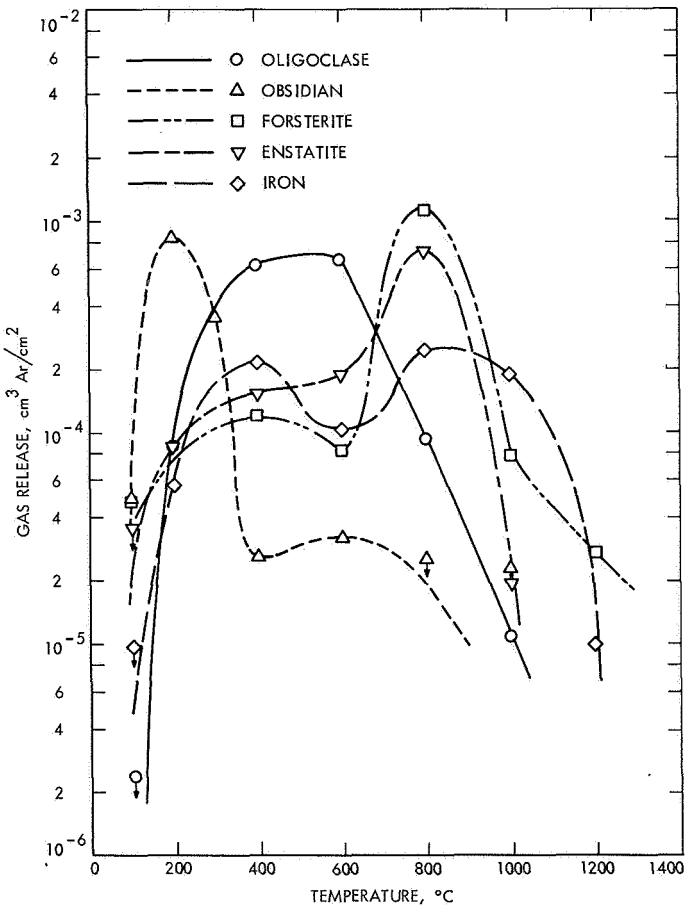


Fig. 7. Differential gas release curves for $2.9 \times 10^{17}/cm^2$ 40-keV argon ions injected into various samples

size, while for the obsidian, enstatite, and forsterite samples, one maximum is significantly larger than the other. The bi-modal distribution can be explained as a low-temperature damage diffusion release which includes point defect diffusion as well as annealing, and a higher temperature normal gas diffusion release.

Pre-irradiation of the samples with 1-keV H^+ (irradiation C, Table 3) resulted in reducing the temperature of maximum argon release for those samples showing a normal gas diffusion release. This decrease in release temperature is consistent with a proton-irradiation-induced metamictization which increases the percentage of the total gas released by damage diffusion. Furthermore, both the forsterite and the iron samples show a more predominant low-temperature release for the higher dose argon irradiation (A) than for the lower dose irradiation (B), due to the same effect.

It is seen from these data that although large incident ion doses disorder the surface increasing the damage diffusion release, the primary control for ion-injected gas release is the substrate composition. The thermal release profiles of solar wind argon ions injected into extraterrestrial material will be reduced to lower temperatures by simultaneous irradiation of the sample by the predominant solar wind ions H^+ and He^{++}. The release of low-energy ion injected gas (as the solar wind) from meteorites, cosmic dust, or lunar surface samples should occur primarily at temperatures below 600° C.

References

1. Lord, H. C., "Hydrogen and Helium Ion Implantation into Olivine and Enstatite," *J. Geophys. Res.*, Vol. 73, pp. 5271–5280, 1968.
2. Davies, J. A., and Jespersgård, P., *Can. J. Phys.*, Vol. 44, 1961.

Erratum

In SPS 37-56, Vol. III, pp. 53–55, the expressions for $Y(\tau';\mu)$ and $z(\tau;\mu)$ should read:

$$Y(\tau';\mu) = Y_0(\tau';\mu) + y(\tau';\mu)$$

$$z(\tau;\mu) = B_\tau\{Y_0(\tau';\mu)\} - \int_0^1 \left[\int_\alpha^\beta Y_0(\tau';\mu)\, X_0(\tau';\mu')\, \Omega(\tau') \exp\{-(\tau-\tau')/\mu'\}\, d\tau'\right] \Psi(\mu')\, d\mu'/\mu'$$

VI. Physics

SPACE SCIENCES DIVISION

A. Relativistic Many-Body Gravity Theory and Celestial Mechanics, *F. B. Estabrook*

General relativistic gravitation theory is now required for analysis of data from ultra-precise radar range tracking of space probes and planets. Conversely, relativistic gravity theory is expected shortly to receive its first really quantitative verification from such range tracking—in particular, from *Mariner VI* and *VII*, Venus–Mercury 1973, and the German–American *Helios* probe. The deviations of the predictions of relativistic theory from those of classical Newtonian theory are nevertheless small on a local, solar-system scale; for large-scale questions of cosmology, and for the long-time history of the universe, galaxy, and planets, relativistic theory has supplied a radically nonclassical world picture, and a very different quantitative understanding.

The two relativistic gravity theories presently acceptable are (1) the tensor theory of Einstein, which is to say, general relativity, and (2) the scalar–tensor theory of Brans and Dicke (Ref. 1), which generalizes the former theory in a formally unique way, involving a long-range scalar field and a theoretically undetermined coupling constant ω. The Brans–Dicke theory reduces to the orthodox Einstein theory in the limit $\omega \to \infty$; if ω were as low as 5, a value not inconsistent with present experimental data, the known perihelion precession of Mercury would in turn imply an anomalously large solar quadrupole moment J_2. Such a J_2 has been supported by recent observations by Dicke's group of the solar optical oblateness. If the apparent oblateness indeed is due to J_2, it would necessitate radical revision of current astrophysical models of the solar interior, and so the solar J_2 has become a matter of high current interest in its own right. In any case, an ω of 5, as emphasized by Dicke, would drastically alter relativistic cosmology and planetology, since the Brans–Dicke theory implies secular change in the universal "constant" of gravitation.

The ultimate verification of relativistic gravity theories, and the accurate evaluation of ω and J_2, will be through celestial mechanical experiments of signal propagation and orbit determination. Much of the analysis of these can be done with single-body perturbation terms calculated solely for the sun itself (using, for example, the Schwarzschild metric of general relativity, or its generalization by Brans–Dicke for finite ω.) When, however, two

or more bodies must be considered in addition to the space probe, the nonlinearity requires a general perturbation analysis. Many-body terms can be required, for example, in the case of lunar theory, or again for the moons of Jupiter, or for precise orbit determination in close encounter with any planet.

The many-body post-Newtonian perturbation terms of the Brans–Dicke theory have now been obtained. The result is a coupled set of simultaneous equations of motion for n gravitating bodies, correct to terms of order $1/c^2$. The equations, like those of the already known limiting case ($\omega \to \infty$, the so-called Einstein–Infeld–Hoffman equations of general relativity), prove to be derivable from a velocity dependent Lagrangian, and seven first integrals of mass, linear momentum, and angular momentum exist (the first four are, surprisingly, independent of ω). The explicit equations, and details of their deviation, are to be published in the open literature.[1] A capability for utilizing them has been added to the JPL double-precision orbit determination program.[2]

Reference

1. Brans, C., and Dicke, R. H., *Phys. Rev.*, Vol. 124, p. 925, 1961.

[1]Estabrook, F. B., "Post-Newtonian n-Body Equations of the Brans-Dicke Theory," *Astrophys. J.*, Vol. 158, Oct. 1969 (in press).

[2]Moyer, T., *Parameterization of Relativity Terms in the DPODP*, Feb. 7, 1969 (JPL internal memorandum).

VII. Spacecraft Telemetry and Command

TELECOMMUNICATIONS DIVISION

A. On the Capacity of a Cascade of Identical Discrete Memoryless Nonsingular Channels,

M. K. Simon

1. Introduction

In practice, one is often concerned with transmitting data over a channel which is composed of a cascade of identical subchannels (e.g., the repeatered telephone line). Calculation of the channel capacity via the standard technique of first finding the overall channel transition matrix becomes extremely tedious as the number of subchannels N becomes large (typically $N = 3000$ for a long-haul telephone link). This article presents an alternative, simplified approach to the problem by combining an eigenvalue technique well known (Ref. 1) in the theory of discrete Markov chains with a channel capacity theorem found in a standard text on information theory (e.g., see theorem 3.3.3 of Ref. 2). The restrictions placed on the subchannels are that they be discrete, memoryless, and have a nonsingular transition matrix.

2. Derivation of the Expression for Channel Capacity

Let $\mathbf{P}_i$ denote the transition matrix of the ith subchannel. Since $\mathbf{P}_i$ is assumed nonsingular, it must be square. If M is the number of messages to be transmitted, i.e., $\mathbf{P}_i$ is $M \times M$, then the input and output state column vectors for the ith subchannel, $\mathbf{p}(Y_{i-1})$ and $\mathbf{p}(Y_i)$, respectively, are related by

$$\mathbf{p}(Y_i) = \mathbf{P}_i^T \mathbf{p}(Y_{i-1}) \tag{1}$$

and $\mathbf{p}(Y_0) \triangleq \mathbf{p}(X)$, the probability distribution vector for the source. For N identical subchannels, the source distribution is related to the output state by

$$\mathbf{p}(X) = [(\mathbf{P}^T)^{-1}]^N \mathbf{p}(Y_N) \tag{2}$$

where the i subscript on $\mathbf{P}^T$ has been dropped since they are all identical. If we let Q denote $\mathbf{P}^{-1}$, it is known (Ref. 1) that $\mathbf{p}(X)$ can be expressed as

$$\mathbf{p}(X) = \sum_{i=1}^{M} a_i \lambda_i^N \mathbf{v}_i \tag{3}$$

where λ_i and $\mathbf{v}_i$ are, respectively, the ith eigenvalue and eigenvector of the matrix $\mathbf{Q}^T$. The coefficients a_i are determined from the set of equations

$$\mathbf{p}(Y_N) = \sum_{i=1}^{M} a_i \mathbf{v}_i \tag{4}$$

Returning now to the computation of channel capacity, we write the mutual information between X and Y_N as

$$I(X;Y_N) = -\sum_{i=1}^{M} p_{yi} \log p_{yi} - \mathbf{H}^T\mathbf{p}(X) \tag{5}$$

where p_{yi} are the elements of $\mathbf{p}(Y_N)$, $\boldsymbol{H}$ is the column vector with ith component

$$H(Y_N|x_i) = -\sum_{j=1}^{M} r_{ji} \log r_{ji} \tag{6}$$

and r_{ij} is the ijth element of the matrix $[\mathbf{P}^T]^N$ (i.e., the transpose of the overall transition matrix). Substituting Eq. (3) in Eq. (5), we obtain

$$I(X;Y_N) = -\sum_{i=1}^{M} p_{yi} \log p_{yi} - \sum_{i=1}^{M} a_i \lambda_i^N \mathbf{H}^T \mathbf{v}_i \tag{7}$$

Letting $u_i = -\mathbf{H}^T\mathbf{v}_i$ and noting that both λ_i and u_i are independent of $\mathbf{p}(Y_N)$, we can maximize $I(X;Y_N)$ with respect to the output distribution using the technique of Lagrange multipliers. Thus, we wish to maximize the function

$$I(X;Y_N) = -\sum_{i=1}^{M} p_{yi} \log p_{yi} + \sum_{i=1}^{M} a_i \lambda_i^N u_i \tag{8}$$

subject to the constraint

$$\sum_{i=1}^{M} p_{yi} = 1 \tag{9}$$

Using the Lagrange multiplier μ, the function to be maximized is then

$$\Phi = -\sum_{i=1}^{M} p_{yi} \log p_{yi} + \sum_{i=1}^{M} a_i \lambda_i^N u_i + \mu\left(\sum_{i=1}^{M} p_{yi} - 1\right) \tag{10}$$

This function will have an extremal value when

$$\frac{\partial \Phi}{\partial p_{yi}} = 0, \qquad i = 1, 2, \cdots, M \tag{11}$$

Thus,

$$\frac{\partial \Phi}{\partial p_{yj}} = \log_2 e\,[-\ln_e p_{yj} - 1] + \sum_{i=1}^{M} \lambda_i^N u_i \frac{\partial a_i}{\partial p_{yj}} + \mu = 0 \tag{12}$$

Solving for p_{yj} gives

$$p_{yj} = e^{-1} \exp_2\left[\mu + \sum_{i=1}^{M} \lambda_i^N u_i \frac{\partial a_i}{\partial p_{yj}}\right] \tag{13}$$

But

$$\sum_{k=1}^{M} p_{yk} = 1 \tag{14}$$

Thus,

$$e^{-1} 2^{\mu} \sum_{k=1}^{M} \exp_2\left[\sum_{i=1}^{M} \lambda_i^N u_i \frac{\partial a_i}{\partial p_{yk}}\right] = 1 \tag{15}$$

Combining Eqs. (13) and (15) gives

$$p_{yj} = \frac{\exp_2\left[\sum_{i=1}^{M} \lambda_i^N u_i \frac{\partial a_i}{\partial p_{yj}}\right]}{\sum_{k=1}^{M} \exp_2\left[\sum_{i=1}^{M} \lambda_i^N u_i \frac{\partial a_i}{\partial p_{yk}}\right]}, \qquad j = 1, 2, \cdots, M \tag{16}$$

Substituting Eq. (16) in Eq. (8) gives[1]

$$C_N = \max_{[Y_N]} I(X;Y_N) = \log_2\left[\sum_{j=1}^{M} \exp_2\left(\sum_{i=1}^{M} \frac{\partial a_i}{\partial p_{yj}} \lambda_i^N u_i\right)\right] \tag{17}$$

By recalling Eq. (4), the partial derivatives $\partial a_i/\partial p_{yj}$ can be evaluated as

$$\frac{\partial a_i}{\partial p_{yj}} = \frac{\Delta_{ji}}{\Delta} \tag{18}$$

where Δ and Δ_{ji} are, respectively, the determinant and cofactor of the jith term of the matrix whose columns are the M eigenvectors. Before proceeding to the evaluation of the partial conditional entropy matrix $\mathbf{H}$ we briefly

[1]The expression for channel capacity as given by Eq. (17) is only valid if there exists an input probability distribution $p(X)$ which gives the output distribution described by Eq. (16).

review some properties of the eigenvalues and eigenvectors of $\mathbf{Q}^T$.

3. Properties of λ_i, ν_i ($i = 1, 2, \cdots, M$)

Since the eigenvalues of $\mathbf{P}^T$ are always less than or equal to 1 in magnitude, the eigenvalues of $\mathbf{Q}^T = [\mathbf{p}^{-1}]^T$, i.e., $\lambda_i, i = 1, 2, \cdots, M$, satisfy

$$|\lambda_i| \geqq 1, \qquad i = 1, 2, \cdots, M \tag{19}$$

Also, $\lambda_i = 1$ is always one of the eigenvalues, and a repeated unity eigenvalue will occur if and only if the matrix $\mathbf{P}^T$ is reducible. The eigenvectors of $\mathbf{Q}^T$ are identical to those of $\mathbf{P}^T$. Thus, the eigenvectors corresponding to the nonunity eigenvalues have components which sum to zero. This property is of use in *Subsection 5*. The elements of the eigenvectors corresponding to the unity eigenvalues are all positive or zero since they must correspond to probabilities.

4. Evaluation of the Matrix H

The kth column of $[\mathbf{P}^T]^N$ denoted by $\mathbf{r}_k$ can be found by post-multiplying $[\mathbf{P}^T]^N$ by $\boldsymbol{\delta}_k$, a column vector with a one in row k and all other elements zero.

Following a procedure similar to that in *Subsection 2*, we write

$$\mathbf{r}_k = \sum_{i=1}^{M} b_i \left(\frac{1}{\lambda_i}\right)^N \boldsymbol{\nu}_i \tag{20}$$

where the b_i satisfy

$$\sum_{i=1}^{M} b_i \boldsymbol{\nu}_i = \boldsymbol{\delta}_k \tag{21}$$

Solving Eq. (21) we get

$$b_i = \frac{\Delta_{ki}}{\Delta} \tag{22}$$

where Δ_{ki} and Δ have the same meaning as previously mentioned. Combining Eqs. (22) and (20) gives

$$\mathbf{r}_k = \sum_{i=1}^{M} \frac{\Delta_{ki}}{\Delta} \left(\frac{1}{\lambda_i}\right)^N \boldsymbol{\nu}_i \tag{23}$$

In terms of Eq. (23), the kth term of the partial conditional entropy matrix may be expressed as

$$H(Y_N|x_k) = -\sum_{j=1}^{M} \left[\sum_{i=1}^{M} \frac{\Delta_{ki}}{\Delta} \left(\frac{1}{\lambda_i}\right)^N \nu_{ji}\right] \times \log\left[\sum_{i=1}^{M} \frac{\Delta_{ki}}{\Delta} \left(\frac{1}{\lambda_i}\right)^N \nu_{ji}\right] \tag{24}$$

where ν_{ji} is the jth component of the ith eigenvector, $\boldsymbol{\nu}_i$. It is at this point that one can observe from Eqs. (17), (18), and (24) and the definition of u_i that the channel capacity is given entirely in terms of the properties of one subchannel transition matrix.

5. Subchannel Transition Matrices Which Are Uniform From the Input and Symmetric

One special case of the above general theory deserves attention. If each subchannel transition matrix is symmetric and uniform from the input, then the overall channel transition matrix will have the same properties. From this, one can conclude that

$$H(Y_N|x_k) \text{ is independent of } k \tag{25}$$

Letting $H(Y_N|x_k) = h$, we can write the partial conditional entropy matrix $\mathbf{H}$ as

$$\mathbf{H} = -h\mathbf{1} \tag{26}$$

where $\mathbf{1}$ is the column vector whose components are all 1.

Recalling the properties of the eigenvectors given in *Subsection 3*, we get

$$u_i = \begin{cases} h \cdot \sum_{k=1}^{M} \nu_{ki}, & \text{all } i \text{ corresponding to eigenvalues 1} \\ 0, & \text{all other } i \end{cases} \tag{27}$$

Thus, Eq. (16) reduces to

$$C_N = \log_2\left[\sum_{j=1}^{M} \exp_2\left(\sum_{i} \frac{\Delta_{ji}}{\Delta} h \sum_{k=1}^{M} \nu_{ki}\right)\right] \tag{28}$$

where the summation on i includes only those terms corresponding to eigenvalues 1. An example of such a subchannel matrix is the binary symmetric matrix (BSC). The

channel capacity for an N-cascade of BSC submatrices is computed in *Subsection 6.*

6. Binary Symmetric Subchannel Matrix

Let

$$\mathbf{P} = \begin{pmatrix} p & q \\ q & p \end{pmatrix} \tag{29}$$

where $q = 1 - p$. Since the matrix $\mathbf{P}$ (or $\mathbf{P}^T$) is irreducible for $0 < p < 1$, only one eigenvalue has value 1. Thus, the summation on i in Eq. (28) will have only one term. The eigenvector corresponding to the eigenvalue 1 is found from the pair of equations

$$\left.\begin{aligned} (p-1)\,v_{11} + q\,v_{21} &= 0 \\ q\,v_{11} + (p-1)\,v_{21} &= 0 \end{aligned}\right\} \tag{30}$$

Thus,

$$v_{21} = -\left(\frac{p-1}{q}\right) v_{11} = v_{11}$$

and normalizing v_{11} to 1, we have

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \tag{31}$$

The other eigenvalue λ_2 has value $1/(p-q)$ and the corresponding normalized eigenvector $\mathbf{v}_2$ becomes

$$\mathbf{v}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \tag{32}$$

The determinant of the eigenvector matrix is found from

$$\Delta = \det \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = 2 \tag{33}$$

and the cofactors Δ_{11} and Δ_{21} are given by

$$\left.\begin{aligned} \Delta_{11} &= 1 \\ \Delta_{21} &= 1 \end{aligned}\right\} \tag{34}$$

The quantity h may be evaluated from Eq. (24) as

$$h = \left[\frac{1}{2} + \frac{1}{2}(p-q)^N\right] \log\left[\frac{1}{2} + \frac{1}{2}(p-q)^N\right] + \left[\frac{1}{2} - \frac{1}{2}(p-q)^N\right] \log\left[\frac{1}{2} - \frac{1}{2}(p-q)^N\right] \tag{35}$$

The channel capacity can now be found by substituting Eqs. (31) through (35) in Eq. (28). The result is

$$\begin{aligned} C_N &= \log_2 [2^{h+1}] \\ &= 1 + \left[\frac{1}{2} + \frac{1}{2}(p-q)^N\right] \log\left[\frac{1}{2} + \frac{1}{2}(p-q)^N\right] \\ &\quad + \left[\frac{1}{2} - \frac{1}{2}(p-q)^N\right] \log\left[\frac{1}{2} - \frac{1}{2}(p-q)^N\right] \end{aligned} \tag{36}$$

The result checks with that given in Problem 3.5 of Ref. 2 obtained from the overall channel transition matrix.

For $N = 1$, we obtain the well-known result

$$C_1 = 1 + p \log p + q \log q \tag{37}$$

References

1. Karlin, S., *A First Course in Stochastic Processes,* pp. 94–107. Academic Press, New York, 1966.
2. Ash, R., *Information Theory,* pp. 56–60. Interscience Publishers, A Division of John Wiley & Sons, Inc., New York, 1965.

B. Efficient Spacecraft Data Systems: Modifications to the *Mariner* Commutator to Adapt to the TOPS Mission, *R. A. Easton and E. E. Hilbert*

The thermoelectric outer-planet spacecraft (TOPS), flight telemetry subsystem (FTS) must handle a large number of sensors (6 to 7 times the *Mariner* requirements) in a flexible formatting arrangement with high reliability over a 12-yr mission (versus 6 mo for *Mariner*). The question arises, can the *Mariner* commutator design be modified to meet these requirements or is a new approach required?

Consider the FTS as a chain with three links (Fig. 1):

(1) The commutator and its driving/decoding programmer circuitry.

(2) The commutation sequence generator [a clock for *Mariner* or a specialized computer in computer-accessed telemetry system (CATS)].

(3) The analog-to-digital converter (A/DC) and signal processing circuitry following the commutator.

It should not be difficult to make link 3 reliable so it will not be discussed here. As for link 2 it should be possible to make the CATS sequence generator highly reliable through the use of error detecting and correcting

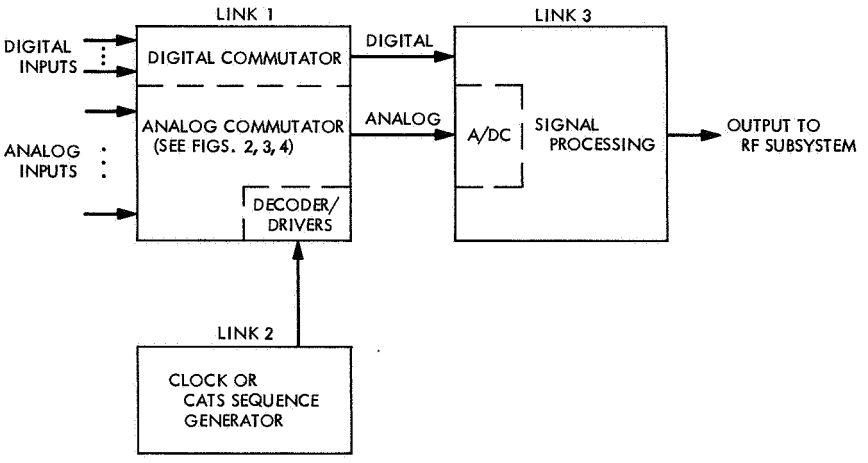


Fig. 1. Flight telemetry subsystem structure

codes. Certainly a simple clock is reliable. This article will concentrate on link 1 and suggest two alternatives to the *Mariner* approach which provide greatly improved reliability and reduce the analog commutator and programmer's power, weight, and volume requirements. The *Mariner* commutator and programmer figures are 2.15 W, 7.5 lb, and 300 in.[3] For TOPS these figures might be 6 W and 15 lb allowed for the entire TOPS FTS; it is clear that some change is needed.

Figure 2 shows the *Mariner* implementation of the commutator and programmer. The clock drives a set of counter flip-flops whose outputs are decoded and fed to drivers each of which drives one of n paralleled field-effect transistor (FET) switches per deck. The FETs are *on* when the drive is removed so a single failure anywhere in the chain resulting in the drive being removed shorts the entire deck. Thus, a large percentage of the possible failures affect a large number of sensors.

Figure 3 shows a modification of the *Mariner* approach for improved reliability. Note that now the FETs are connected in a self-decoding tree structure so that fewer drivers and no decoding gates are required. Also note that most failures now affect much fewer sensors. This clocked tree approach can be used as a 1-to-1 replacement for the *Mariner* subdecks, since it generates its own fixed sequence as does the *Mariner* approach and requires only a clock for a driver.

The third approach, for use with a programmable sequence generator such as CATS, is to use a large random access tree for the whole system instead of using several smaller trees as subdecks. There is no subcommutation required within the commutator because the sequence is loaded from a computer memory (Fig. 4).

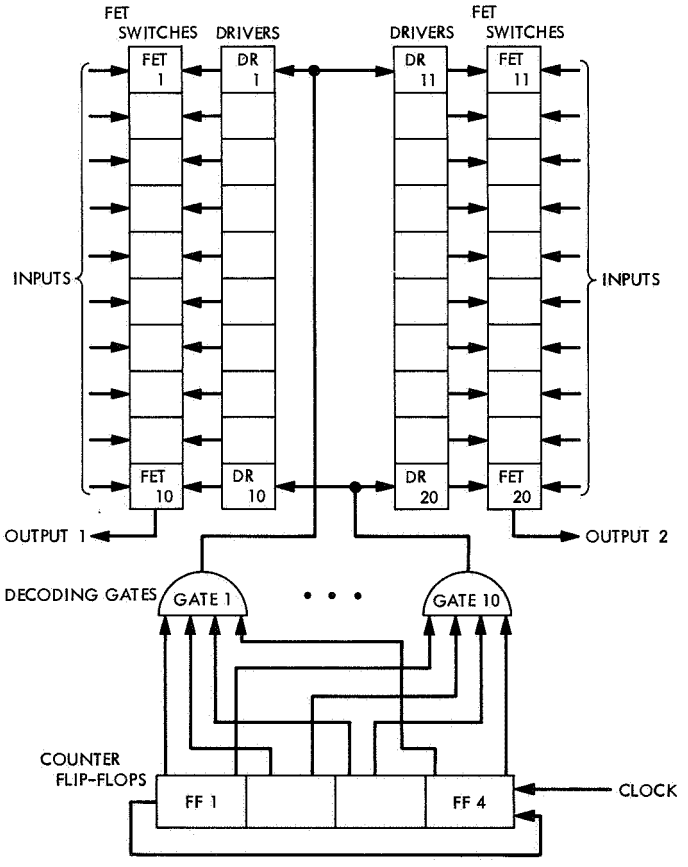


Fig. 2. *Mariner* commutator (two decks shown)

Figure 5 compares the reliability of these three approaches for the *Mariner* Mars 1969 sensor requirements

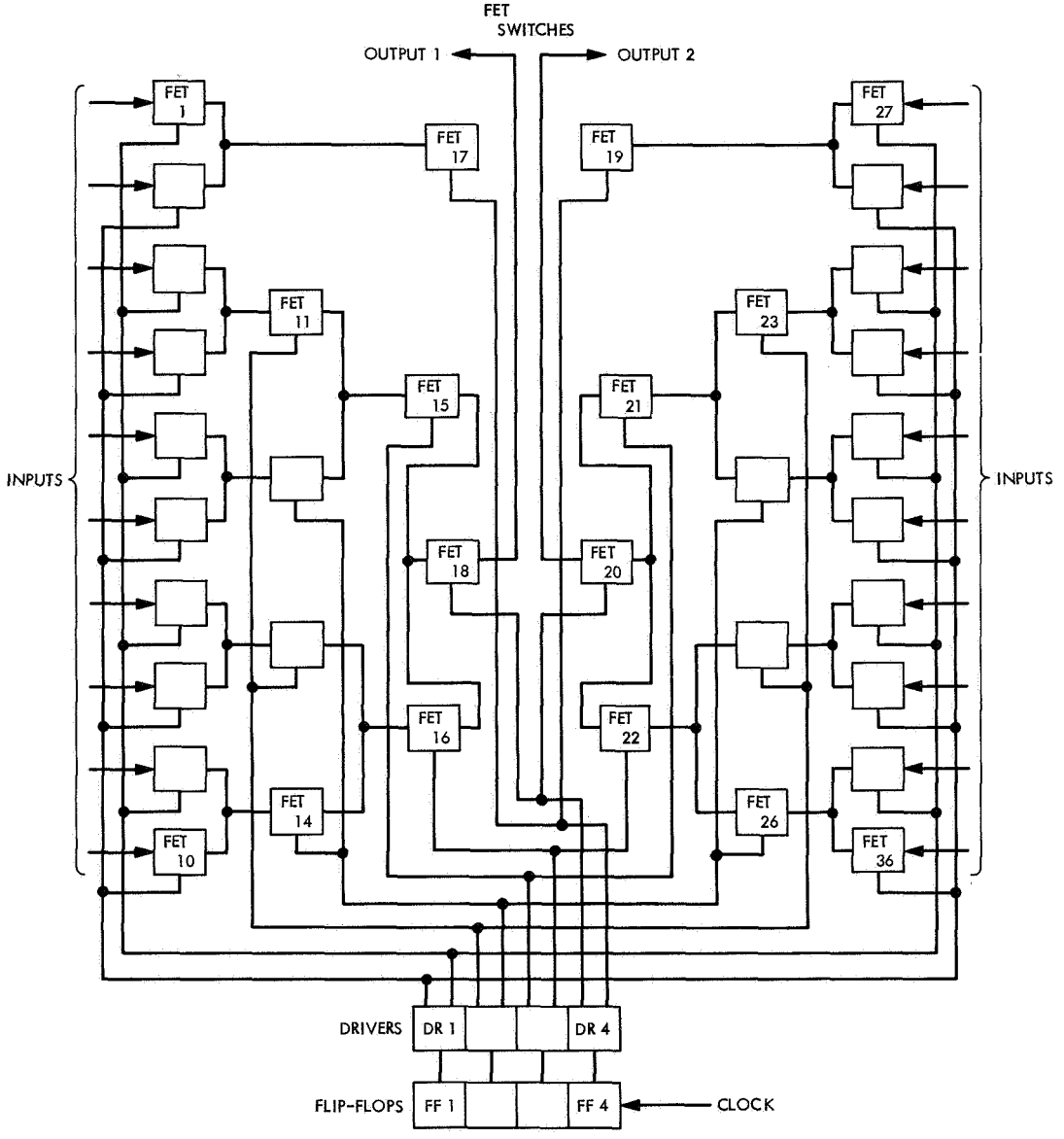


Fig. 3. Tree-type commutator (two decks shown)

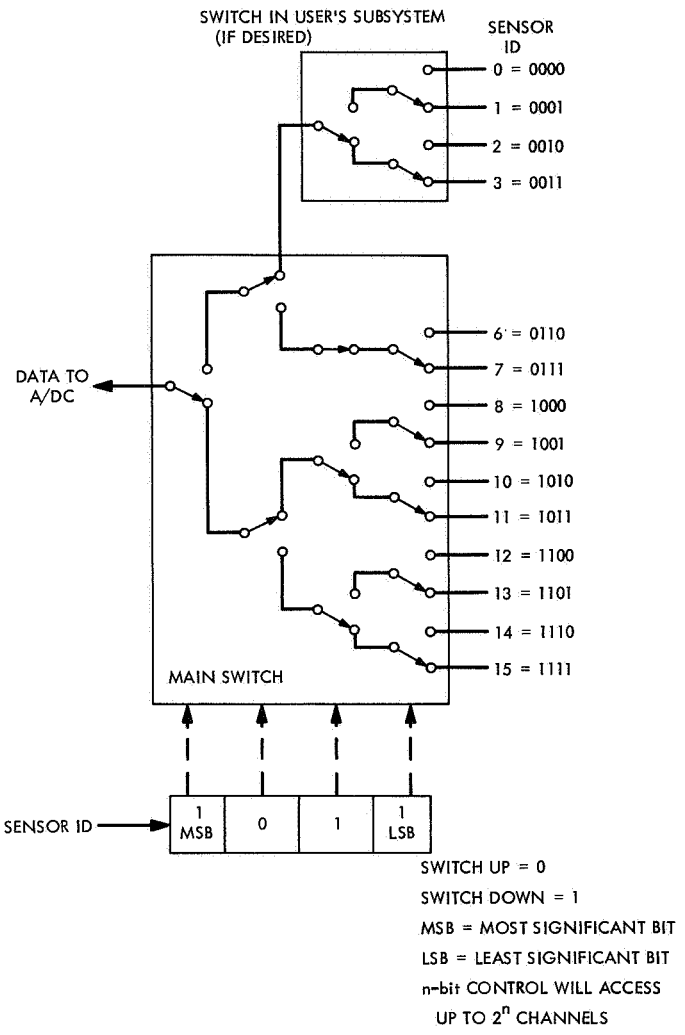


Fig. 4. Random access tree switch

as a function of time. The failure rates assumed are shown in Table 1 and are realistic numbers for space applications. For a Mars mission the reliability of the *Mariner* system looks satisfactory, but even for this small number of sensors the likelihood for unimpaired operation after 12 yr (a 10.5% chance) is poor. The reliability guideline for the TOPS project is 95% for a 12-yr mission for important subsystems. With five times as many sensors, the prospects look grim indeed. The tree's chances look much better but with no redundancy some failures might still occur. Compared to the *Mariner* approach, only a small percentage of the parts involved (i.e., the drivers and a few switches common to a large number of paths) are in a position to affect a large number of sensors if they fail. However, it is possible to make these elements highly reliable by making the drivers parallel redundant and the tree elements serial/parallel redundant. Implemented in

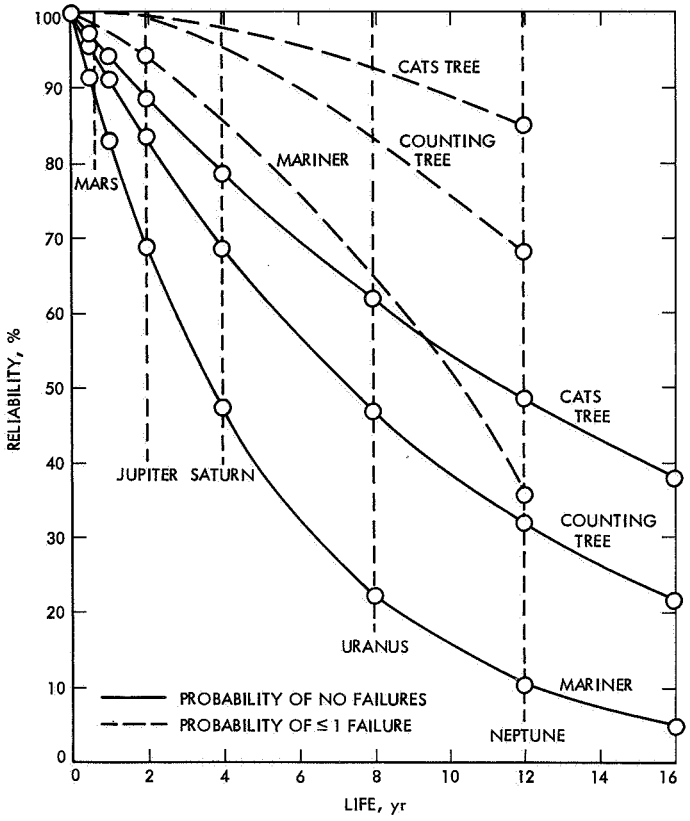


Fig. 5. Comparison of reliability as a function of time

this way, no single component failure can degrade the system operation and a large number of random failures can be tolerated without significant degradation.

This is not the case for the *Mariner* approach because, in order to get a substantial improvement in reliability, the decoding gates and most of the drivers must be made redundant. There are nearly as many drivers as there are FETs. The decoding gates cannot be made parallel redundant without power switching since their outputs would have to be *or*ed together. This switching would be at

Table 1. Failure rates used in deriving reliability curves[a]

Component	Number of failures/h
Resistors	0.001×10^{-6}
Capacitors	0.005×10^{-6}
Diodes	0.01×10^{-6}
Transistors	0.02×10^{-6}
FETs	0.02×10^{-6}
Monolithic ICs	0.20×10^{-6}
Mariner driver (discrete)	0.10×10^{-6}
CATS driver	0.20×10^{-6}

[a] Assume failures leading to FETs staying *on* are equally likely as failures leading to FETs staying *off*.

least as complicated as the gates themselves, limiting substantially the improvement obtainable.

A major advantage of the tree approach is the ease with which redundancy can be implemented. As the tree grows in size, this condition gets even better since the percentage of parts that will affect a given percentage of sensors steadily decreases because only n drivers are required for 2^n sensors and a fixed percentage of switches affect a fixed number of sensors. Therefore, if the number of sensors doubles, the percentage affected is halved.

In Fig. 6 the probabilities of $\leq$1 failure and no failures and the average number of sensors lost in the first failure are plotted for the three systems as a function of the number of flatpacks required to implement the systems with presently available components. The figures given are for a 12-yr mission. The three points on each curve represent (from left to right) three options:

(1) No redundancy.

(2) Redundant flip-flops and decoding gates for *Mariner*; redundant flip-flops and drivers for the trees.

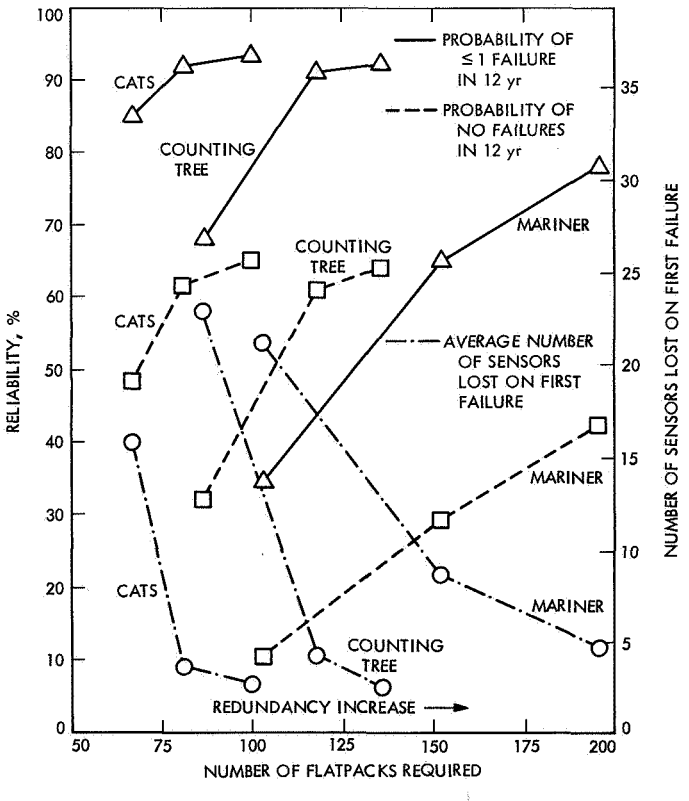


Fig. 6. Comparison of reliability as a function of complexity

(3) The redundancy of option 2 plus serial/parallel redundancy of all FETs (and corresponding drivers for the *Mariner* case) that could affect more than 10 of the 77 sensors.

Note how fast the redundancy takes effect for the trees as compared to the *Mariner* system. In fact, it takes fewer parts to implement the third level of redundancy on the CATS (random access) tree than it takes for a nonredundant *Mariner* system. The comparison between these two points is given in Table 2.

Table 2. Comparison of *Mariner* and CATS (random access) tree commutators using same number of flatpacks

Parameter	*Mariner* Mars 1969	CATS
Probability of no failure, %	10.6	65.1
Probability of $\leq$1 failure, %	36.7	93.1
Probability of 2 or more failures, %	63.3	6.9
Average number of sensors lost on first failure	21.5	2.8

Using paralleled flip-flop/driver circuits driving a CATS tree whose elements are serial/parallel redundant, it should be possible to make a telemetry system with 500 sensors with a 99% chance of no failures over a 12-yr mission. Referring to Fig. 7, realistically it would be more sensible to allow the last 3 rows of the FET switch tree to be nonredundant, saving 2700 FETs and giving over 99% chance of no failures in the rest of the system and for the nonredundant portion:

$$P(\text{no failures}) = 15.2\%$$

$$P(\leq 1 \text{ failure}) = 44.1\%$$

$$P(\leq 2 \text{ failures}) = 71.2\%$$

$$P(\leq 3 \text{ failures}) = 88.2\%$$

$$P(\leq 5 \text{ failures}) = 99.2\%$$

at 1.7 sensors lost/failure.

Even if 3 FETs failed, fewer than 1% of the sensors would be lost on the average after 12 yr (this corresponds to the right-hand point on the CATS curves in Fig. 6). A 1% loss need not be serious because many of the TOPS sensors are redundant and for critical sensors it is easy to provide alternate paths that have no failures in common by assigning the complement of the original address to the redundant path. This introduces no timing problems due to the random access nature of the tree and no single

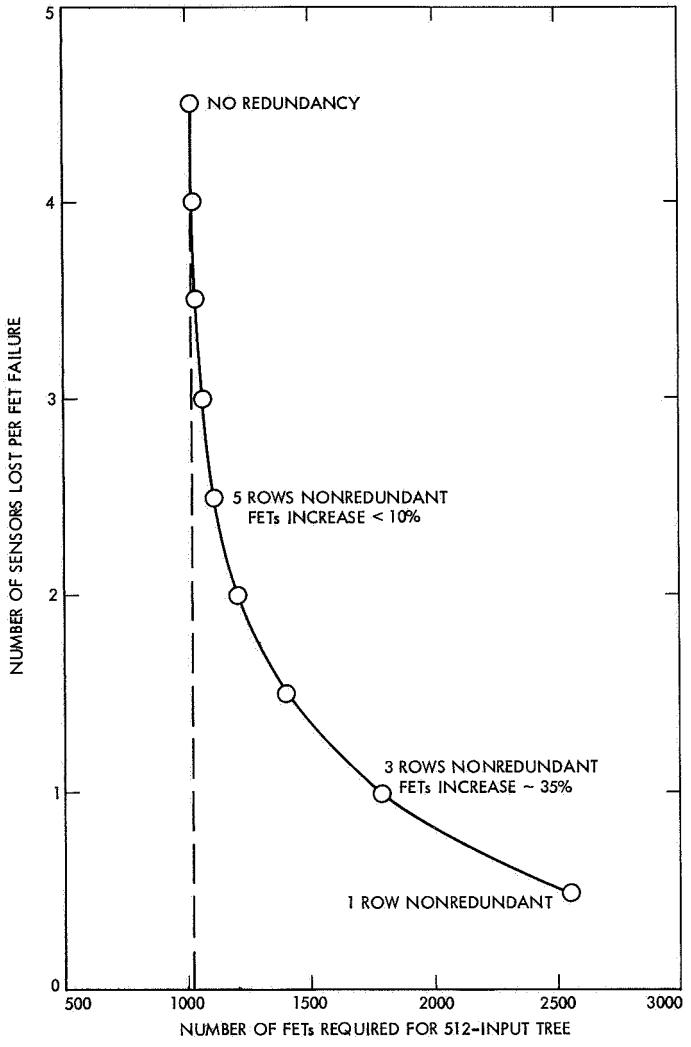


Fig. 7. Tree reliability versus redundancy

failure can cause both paths to fail. Another advantage of this method is the fact that some of the tree branches can be placed in remote locations in the user's subsystem to reduce the number of wires required between subsystems. This would not be practical in the clock-driven systems where the subsystem's measurements are on several different subdecks. In TOPS the power, attitude control, and radio subsystems are possible candidates for remote tree branches.

It is felt that a strong case can be made for switching to the tree commutator approach. The CATS (random access) tree is the most reliable and flexible but if only a clock is available for sequencing (as in the *Viking* project) the clocked tree can provide a substantial reliability improvement over the *Mariner* approach along with a reduction in power, weight, volume, and components. A comparison of the present *Mariner* Mars 1969 commutator and programmer with a random access tree commutator and programmer meeting TOPS requirements is made in Table 3.

Table 3. Comparison of the present *Mariner* Mars 1969 analog commutator and programmer with a random access tree commutator and programmer meeting TOPS requirements

Parameter	*Mariner* Mars 1969	TOPS
Number of analog sensors	77	512
Power, mW	2150	20
Weight, lb	7.5	2.0
Volume, $in.^3$	300	100
Probability of surviving mission with less than 1% channel loss	0.91 after 6 mo	0.99+ after 12 yr

VIII. Spacecraft Telecommunications Systems

TELECOMMUNICATIONS DIVISION

A. Frequency Tracking in an MFSK Receiver,

H. D. Chadwick

1. Introduction

This is one of a series of articles (SPS 37-48, Vol. III, pp. 252–264 and SPS 37-52, Vol. III, pp. 239–247) treating the problems of synchronization of an m-ary frequency shift keying (MFSK) receiver in time and frequency.[1] The previous works are concerned with time and frequency acquisition and time tracking. In this article, a frequency tracking technique is proposed and analyzed which permits tracking of long-term frequency drift with a resultant detection performance loss on the order of 1 dB.

2. Frequency Tracking Problem

The frequency instability of the received MFSK signal has been separated, for purposes of analysis, into long-term and short-term components. The long-term component contains that part of the drift whose frequency spectrum lies below $1/T$, where T is the signaling interval and the integration time of the receiver. Figure 1 illustrates a possible frequency drift pattern and shows the division into long- and short-term components. Only the long-term component is considered here, with correction performed by a digital tracking loop in the receiver. Compensation for short-term drift requires a modification of the detection technique itself.[2,3] It is, therefore, assumed here that short-term drift does not exist to the extent that it affects the design of a long-term drift tracking system, and that the detector uses the direct Fourier transform-power spectrum technique described in previous articles.

Figure 2 illustrates the basic MFSK receiver. The receiver bases its decision on an estimated Fourier power spectrum of the received input. The probability distribution of the spectral lines has been derived for the case of a fixed frequency offset, $\Delta f = \rho/T$, from the nominal frequency value k/T. This frequency offset is a function of the received frequency and the local oscillator setting, and it can be seen from Fig. 2 that the offset may be reduced to zero by a correct setting of the local oscillator.

[1]Springett, J. C., *A Synopsis on Low Data Rate Direct Link Requirements and Technology*, 1969 (JPL internal document) covers in detail the systematic nature of the synchronization problem.

[2]Gleicher, N., *Incoherent Detection of Low Data Rate MFSK with Noisy Oscillators*, Dec. 1968 (JPL internal document).

[3]Gleicher, N., *Semicoherent Detection of Low Data Rate MFSK with Noisy Oscillators*, Dec. 1968 (JPL internal document).

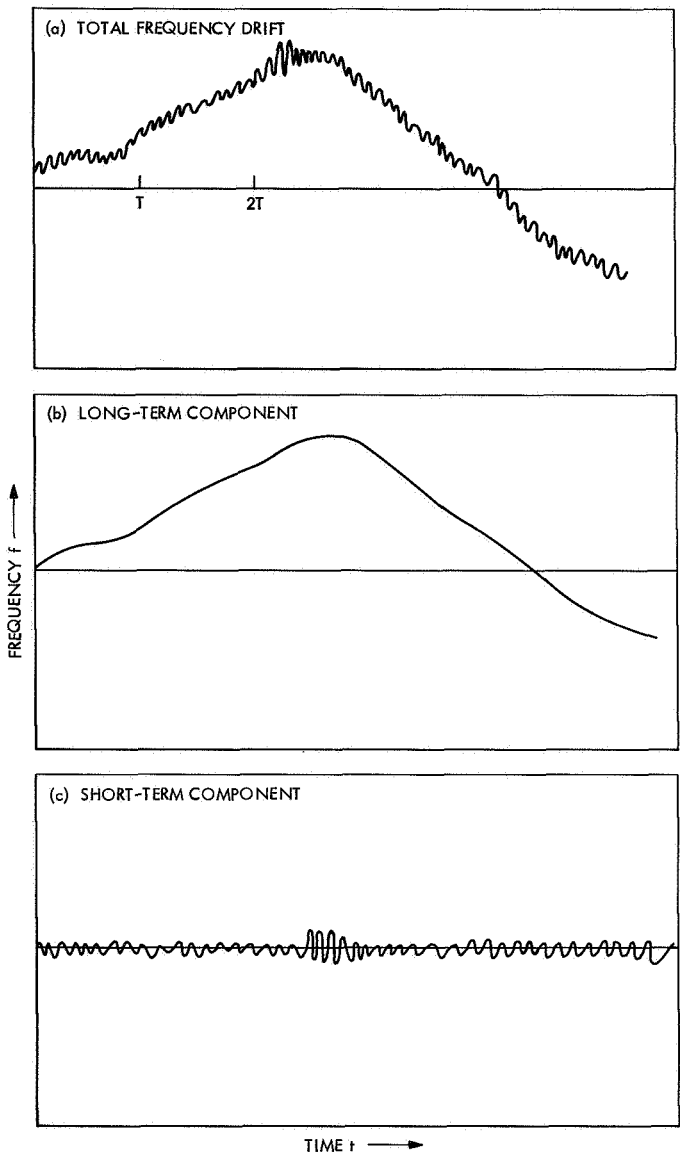


Fig. 1. Frequency drift pattern with long- and short-term components

The performance of the MFSK detector is best when $\Delta f = 0$, or the demodulated signals at the input to the detector occur at frequencies which are integer multiples of $1/T$. The objective of the frequency tracking loop is, therefore, to minimize, in some sense, the average frequency error Δf.

3. Possible Approaches

Two approaches to the frequency tracking problem have been suggested, by Ferguson (Ref. 1) and Goldstein (SPS 37-55, Vol. II, pp. 30–36). Both techniques involve deriving an estimate of the frequency error from the magnitudes of the spectral components neighboring the nominal signal frequency.[4] Ferguson's technique uses a weighted average of the n closest spectral components, where n is some small integer. Mathematically, if the received signal is nominally at frequency k/T, and the corresponding spectral component is r_k, the estimate is given by

$$\hat{\Delta f} = C \sum_{i=-n/2}^{n/2} a_i \, r_{k+i}$$

where C is a normalizing constant and the a_i's are a set of linearizing weighting coefficients.

Goldstein's technique uses a nominal signal frequency at $(k + 1/2)/T$, centered between two spectral lines, and uses the difference between the magnitudes of the two neighboring lines, or

$$\hat{\Delta f} = r_{k+1} - r_k$$

In the following subsection, several different forms of the frequency error estimator are considered.

4. Frequency Error Estimator

The estimation of the frequency error Δf is a statistical problem. The estimator $\hat{\Delta f}$ must be a function of the random variables r_k and ideally should have the characteristics

$$E(\hat{\Delta f}) = \Delta f$$

and

$$\operatorname{var} \hat{\Delta f} = E(\hat{\Delta f}^2) - E^2(\hat{\Delta f}) = \text{minimum}$$

over the class of all possible estimates.

The probability density function of the spectral components r_k has been shown to be

$$p(r_k) = \frac{r_k}{\sigma^2} \exp\left\{-\frac{1}{2\sigma^2}\left[r_k^2 + B^2 \operatorname{sinc}^2 \pi (k - j - \rho)\right]\right\} \times I_0\left[\frac{r_k B}{\sigma^2} \left|\operatorname{sinc} \pi (i - \rho)\right|\right] \quad (1)$$

[4]It is always assumed that the nominal signaling frequency is known, i.e., that the frequency error Δf is less than half the signal spacing, and that the detector is capable of making a reliable decision with this amount of error. This will be shown later to be a good assumption.

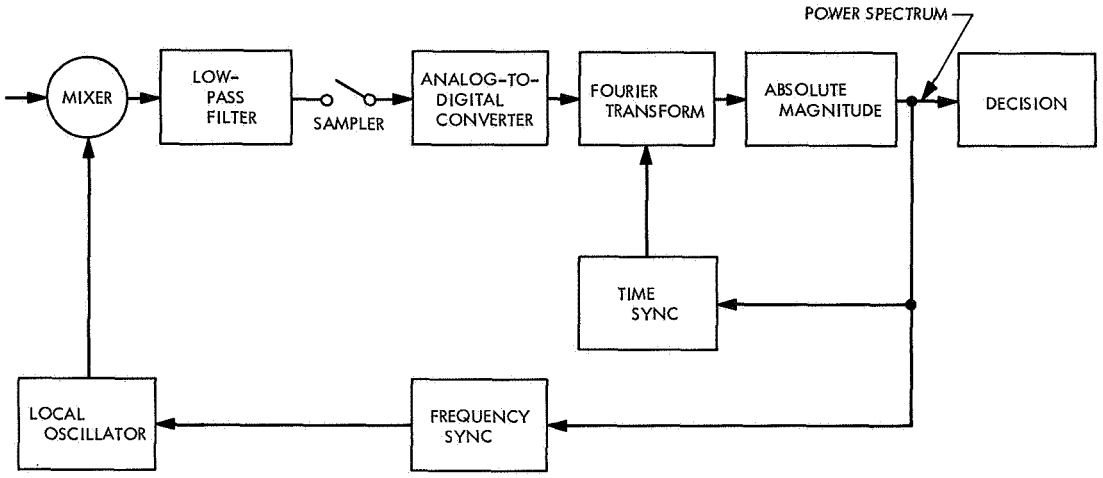


Fig. 2. MSFK receiver block diagram

where

$\sigma^2 = (N_0/4T)\,N^2$

$B = AN/2$

T = signal duration

A = signal amplitude

$N = 2WT$ = number of samples in T

$N_0/2$ = two-sided noise spectral density[5]

k/T = observed frequency

j/T = nominal signaling frequency

$\Delta f = \rho/T$ = frequency error

and

$$\operatorname{sinc} x = \frac{\sin x}{x}$$

Concentrating on the components nearest to the signal frequency ($k \simeq j$) and replacing the index k by $i = k - j$ yields

$$p(r_i) = \frac{r_i}{\sigma^2} \exp\left\{-\frac{1}{2\sigma^2}\left[r_i^2 + B^2 \operatorname{sinc}^2 \pi(i-\rho)\right]\right\} \times I_0 \frac{[r_i B \operatorname{sinc} \pi(i-\rho)]}{\sigma^2} \tag{2}$$

Obviously, all of the r_i terms contain information about the parameter ρ, but as $|i|$ increases, the $\operatorname{sinc}^2 \pi(i-\rho)$ term decreases. The probability densities of the spectral lines corresponding to frequencies far away from the signal frequency are only slightly affected by ρ. Only the line at frequency $i = +1$ is very much affected when ρ is greater than zero, and correspondingly only the line at $i = -1$ is much affected when $\rho < 0$.[6] Thus, it is reasonable to use the values of these nearby lines when constructing an estimator for ρ. Using lines further away will only increase the information gained very slightly while greatly increasing the variance of the estimate.

Two approximations can be made to the maximum likelihood estimator, based on the three spectral lines, r_{-1}, r_0, and r_{+1}. These estimators are

$$\hat{\rho}_1 = \frac{r_{+1} - r_{-1}}{r_{+1} + r_{-1}} \tag{3}$$

and

$$\hat{\rho}_2 = \frac{r_{+1} - r_{-1}}{2r_0} \tag{4}$$

A third estimator is also considered here on the grounds that it is less biased than $\hat{\rho}_2$. (Its expected value is closer to ρ over the range $-1 < \rho < 1$.) This estimator is given by the relation

$$\hat{\rho}_3 = [\operatorname{sgn}(r_{+1} - r_{-1})]\left(\frac{r_{+1} + r_{-1}}{r_0}\right) \tag{5}$$

[5]To conform to common usage, $N_0/2$ is used here instead of N_0 which was used in previous articles. For this reason, the curves in SPS 37-48, Vol. III, pp. 252–264, and SPS 37-52, Vol. III, pp. 239–247, are 3 dB higher than here.

[6]This effect is illustrated in Figs. 2 through 4 of SPS 37-52, Vol. III, pp. 239–247.

where

$$\operatorname{sgn} x = \begin{cases} +1 & x \geqq 0 \\ -1 & x < 0 \end{cases}$$

In order to choose between these different estimators, it is necessary to determine their means and variances as a function of ρ. This calculation is performed in *Subsection 5.*

It may be noticed that none of these estimators are of the form used by Ferguson or Goldstein. All three tentative estimators are ratios between the magnitude of the spectral lines. This tends to reduce the effect of varying signal and noise levels on the magnitude of the estimate. The staggered technique proposed by Goldstein was not included because it requires either a loss in the detection process of about 3 dB, or the computation of a second Fourier transform during each decision interval at the displaced frequency.

5. Calculation of the Probability Density Function (PDF) of the Estimators

The choice between the tentative estimators is made on the basis of: (1) their degree of bias, or how closely $E(\hat{\rho})$ agrees with ρ, and (2) their comparative variance under the same noise and signal conditions. The means and variances are computed by deriving expressions for the probability density function and integrating.

The resulting PDFs, none of which can be found in closed form, are found by integrating the following expressions numerically;[7] simultaneously, the means and variance can be determined:

$$p(\hat{\rho}_1) = \frac{2(1-\hat{\rho}_1)}{(1+\hat{\rho}_1)^3} \int_0^\infty u^3 \exp\left(-\frac{1}{2}\left\{u^2 + \frac{u^2(1-\hat{\rho}_1)^2}{(1+\hat{\rho}_1)^2} + \gamma^2\left[\operatorname{sinc}^2 \pi(1-\rho) + \operatorname{sinc}^2 \pi(1+\rho)\right]\right\}\right)$$
$$\times I_0\left[u\gamma\left|\operatorname{sinc} \pi(1-\rho)\right|\right] I_0\left[\frac{u(1-\hat{\rho}_1)^2}{(1+\hat{\rho}_1)}\gamma\left|\operatorname{sinc} \pi(1+\rho)\right|\right] du \tag{6}$$

$$p(\hat{\rho}_2) = \iint_R \frac{uv(u-v)^2}{\hat{\rho}_2^3} \exp\left(-\frac{1}{2}\left\{u^2 + v^2 + \left(\frac{u-v}{\hat{\rho}^2}\right)^2 + \gamma^2\left[\operatorname{sinc}^2 \pi\rho + \operatorname{sinc}^2 \pi(1-\rho) + \operatorname{sinc}^2 \pi(1+\rho)\right]\right\}\right)$$
$$\times I_0\left[u\gamma\left|\operatorname{sinc}^2 \pi(1+\rho)\right|\right] I_0\left(\frac{u-v}{\rho_2}\gamma\left|\operatorname{sinc} \pi\rho\right|\right) I_0\left[v\gamma\left|\operatorname{sinc} \pi(1-\rho)\right|\right] du\, dv \tag{7}$$

$$p(\hat{\rho}_3) = \iint_R \frac{uv(u+v)^2}{\rho_3^3} \exp\left(-\frac{1}{2}\left\{u^2 + v^2 + \left(\frac{u+v}{\hat{\rho}_3}\right)^2 + \gamma^2\left[\operatorname{sinc}^2 \pi\rho + \operatorname{sinc}^2 \pi(1-\rho) + \operatorname{sinc}^2 \pi(1+\rho)\right]\right\}\right)$$
$$\times I_0\left[u\gamma\left|\operatorname{sinc} \pi(1+\rho)\right|\right] I_0\left(\frac{u+v}{\hat{\rho}_3}\gamma\left|\operatorname{sinc} \pi\rho\right|\right) I_0\left[v\gamma\left|\operatorname{sinc} \pi(1-\rho)\right|\right] du\, dv, \qquad \gamma^2 = \frac{2E}{N_0} \tag{8}$$

where R is the region defined by

$$R = \begin{cases} (u, v : u < v) & \text{for } \hat{\rho} < 0 \\ (u, v : u > v) & \text{for } \hat{\rho} > 0 \end{cases}$$

6. Choice of Estimator

The variances of the three different estimators over the typical range of E/N_0 (signal energy to noise spectral density ratio) are shown in Fig. 3. The means are plotted in Fig. 4 versus the true value of normalized frequency error ρ. It may be seen that while $\hat{\rho}_1$ gives the best performance in terms of linear response to changes in ρ, the estimator $\hat{\rho}_2$ has the lowest variance. For this reason, $\hat{\rho}_2$ was chosen as the proper estimator to use for frequency tracking. It was felt that nonlinearity of the mean $E(\hat{\rho})$ may be more easily corrected than an excessively high

[7] A more complete derivation is presented in Chadwick, H., *Frequency Tracking in an MFSK Receiver*, Apr. 1969 (JPL internal document).

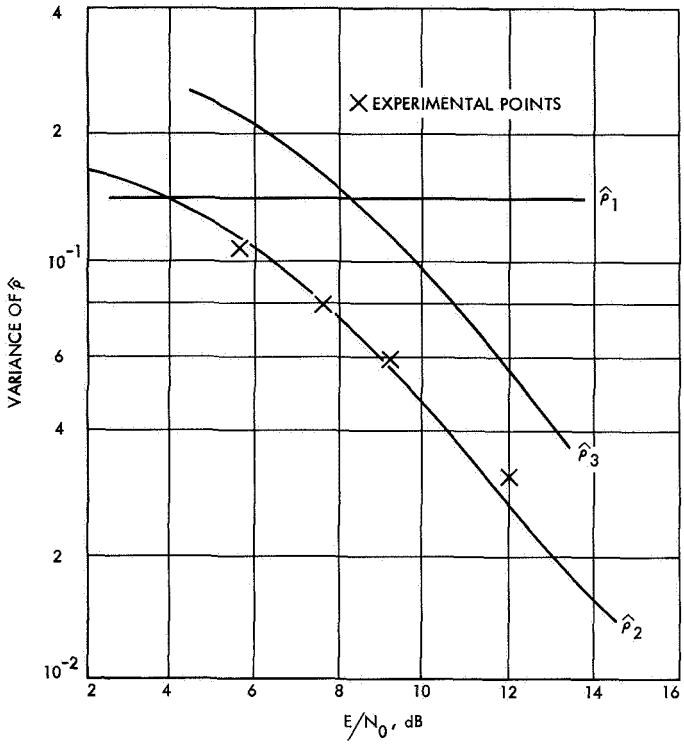


Fig. 3. Variance of estimator $\hat{\rho}$ vs signal-to-noise ratio ($\rho = 0$)

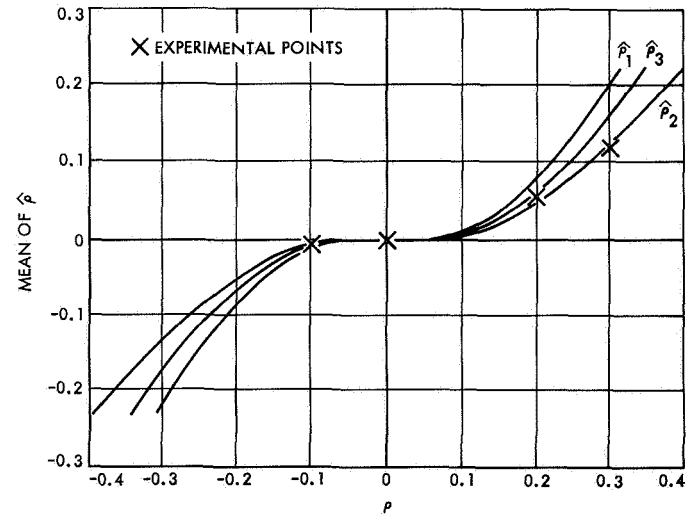


Fig. 4. Mean value of estimator $\hat{\rho}$ vs actual value ρ

variance, which would require a high degree of filtering and a correspondingly lower response time to changes in the frequency error. Experimental results for the statistics of the estimator $\hat{\rho}_2$ are also shown in Figs. 3 and 4. These results were obtained by using an MFSK simulation program on a digital computer, and obtaining the estimates without closing the frequency tracking loop. The experimental results agree very closely to the predicted results.

7. The Frequency Tracking Loop

With the estimator $\hat{\rho}_2$, which will henceforth be called $\hat{\rho}$, it is now necessary to design the frequency tracking loop. The most general form of this loop is shown in Fig. 5a. The output of the estimator box, $\hat{\rho}$, may be separated into two terms

$$\hat{\rho} = E(\hat{\rho}) + N_\rho = g(\rho) + N_\rho \tag{9}$$

where N_ρ is a zero mean random variable with distribution given by Eq. (7). N_ρ may be treated as an additive noise term and the function $g(\rho) = E(\hat{\rho})$ as a nonlinear operation on the time error ρ produced by the estimator. The resulting block diagram is Fig. 5b. By making one more substitution, $F(s) = SH(s)$, the block diagram is

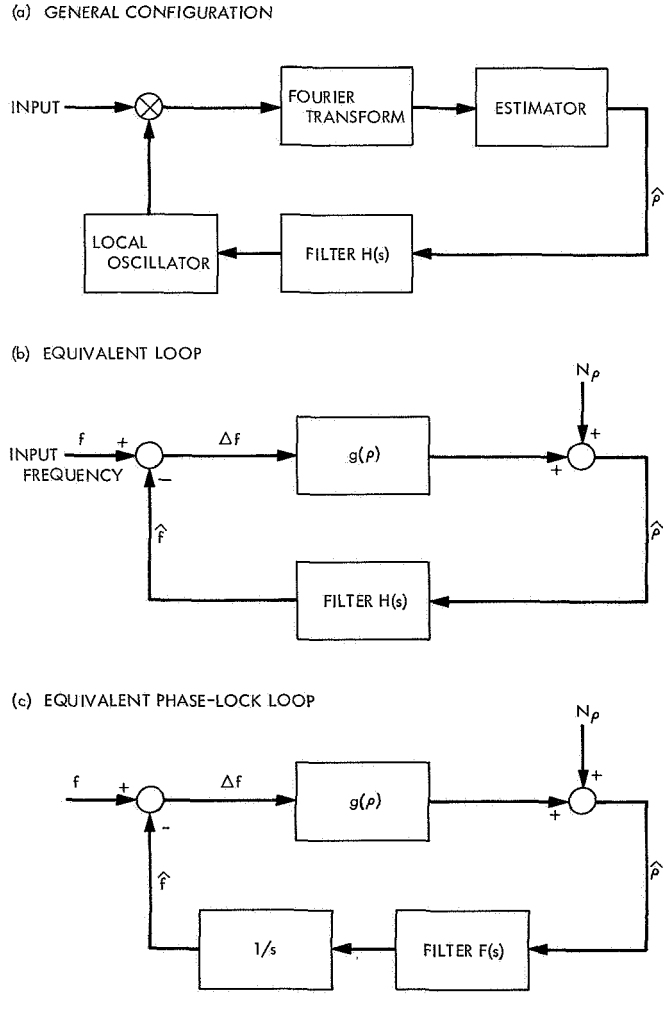


Fig. 5. Frequency tracking loop

changed to Fig. 5c, and it may be seen that this is equivalent to the standard block diagram of a phase-lock loop, with nonlinearity $g(\rho)$ replacing that of the phase detector.

The curve of $E(\hat{\rho})$ versus ρ was found to be approximated closely by the function

$$E(\hat{\rho}) = 5\rho^3 \tag{10}$$

The closeness of this approximation for $p \leq 0.3$ may be seen from Fig. 6. Thus, the nonlinear function $g(\rho)$ can be replaced by this cubic expression. The obvious disadvantage of such a function is its zero slope through the origin. An ideal function would have a slope of one at the origin. However, the rapidly increasing slope of $g(\rho)$ for $\rho \neq 0$ is a good quality for an error signal to possess.

8. Tracking Loop Filter Design

The highly nonlinear nature of the function $g(\rho)$ makes design of the tracking loop filter difficult by traditional linear methods. The design criterion chosen was that the mean square frequency error $E(\Delta f)^2$ is less than 0.0004. This is dictated by the fact that, for the design parameters used in the experimental program ($T = 10$ s), $\rho = 0.2$ corresponds to $\Delta f = 0.02$. This value of ρ results in an average performance loss of about 1 dB. An rms error of 0.02 was thus felt to be acceptable. The input frequency error was assumed to be linearly increasing with a maximum slope of 0.0025 Hz/s. A design point of $E/N_0 = 12$ dB was chosen because the probability of word error at this point is approximately 0.01, which was considered the lowest operating point of an MFSK system.

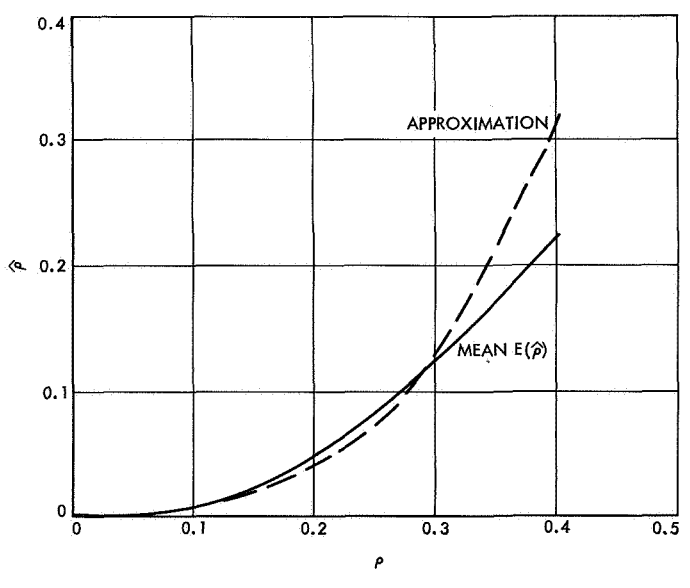


Fig. 6. Approximation to mean $E(\hat{\rho})$ at $E/N_0 = 12$ dB

With these design goals, a second-order loop was synthesized and its theoretical performance calculated by two different methods: (1) the use of a linear approximation to the nonlinearity $g(\rho)$ with the standard linear phase-lock loop approximation; and (2) the technique due to Lindsey (Ref. 2) for nonlinear phase-lock loop analysis. The substitution $F(s) = SH(s)$ was performed to convert the loop to an equivalent phase-lock loop.

Taking the linear phase-lock loop results in Viterbi (Ref. 3, p. 35) that

$$\sigma^2_{\Delta f} = \frac{N_0 W_L}{2A^2} \tag{11}$$

and substituting $N_0 = \sigma_N^2 T$, where σ_N^2 is the variance of the estimator $\hat{\rho}/T$ at the design point

$$\frac{E}{N_0} = 12\,\text{dB}$$

and A is the linearized gain of the estimate, gives the approximate result

$$W_L = 10^{-2}\,\text{Hz}$$

The linearized gain was formed by drawing a straight line through the $g(\rho)$ curve at the points $\rho = \pm 0.2$. The slope of this line at $E/N_0 = 12$ dB is 0.25. The filter was assumed to be of the form

$$F(s) = \frac{C_1}{s} + C_2$$

and choosing the damping factor to be $2^{1/2}$ (typical to phase-lock loop practice) and the loop gain K to be 1 gives the solution

$$C_1 = 3.6 \times 10^{-4}$$

$$C_2 = 5.4 \times 10^{-2}$$

The value of the variance of the frequency error, predicted by the linear approximation using these constants, is

$$\sigma^2_{\Delta f} = 2.6 \times 10^{-4}$$

With Lindsey's technique (Ref. 2), a more accurate prediction of the variance of the frequency error is possible. A difficulty arises in applying the linearized solution he derives when the slope of the nonlinearity $g(\rho)$ is zero at the origin, since his derivation assumes that $g(\rho)$ is normalized to have unit slope at the origin and that the gain is included in the constant A. This problem arises only

when applying the results in Lindsey's paper which deal with a specific solution for a second-order phase-lock loop. The basic results of his paper, however, are independent of the form of $g(\rho)$, and it is presumably possible to derive a specific solution for the special case $g'(0) = 0$. This detail is not attempted here since the results appear to be sufficiently accurate without considering this problem. Since it is impossible to normalize the slope $g'(0)$ to unity, there is some question as to what the value of A should be. It was decided to let $A = 1$ in this case (since physically there is no gain in the loop at this point) and to determine the predicted $\sigma^2_{\Delta f}$ on this basis. Computation of $\sigma^2_{\Delta f}$ requires the simultaneous solution of the two equations

$$\alpha = -\left[\left(\frac{r+1}{r}\right)R - \frac{1}{r\sigma^2_G}\right] \tag{12}$$

and

$$\sigma^2_G = \int_{-\infty}^{\infty} [g(\rho)]^2 p(\rho)\, d\rho \tag{13}$$

where

$$p(\Delta f) = C_0 \exp\left[-\alpha \int^{\Delta f} g(x)\, dx\right] \tag{14}$$

and

$$r = 2, \qquad R = \frac{2A^2}{\sigma^2_N T W_L}$$

From the solution, $p(\Delta f)$ is found and its variance may then be computed.

The solution was determined graphically as illustrated in Fig. 7. The intersection of the lines due to the two equations gives the solution point, which is, in this case

$$\alpha = 5700$$

$$\sigma^2_G = 1.2 \times 10^{-4}$$

The illustration also shows that

$$\sigma^2_{\Delta f} = 3.7 \times 10^{-4}$$

A second solution with A assumed to be 0.25 is also shown in Fig. 7 and gives the result

$$\sigma^2_{\Delta f} = 4.4 \times 10^{-4}$$

Since all results are on the same order as the desired range, it was decided to mechanize the loop, and to determine its behavior experimentally.

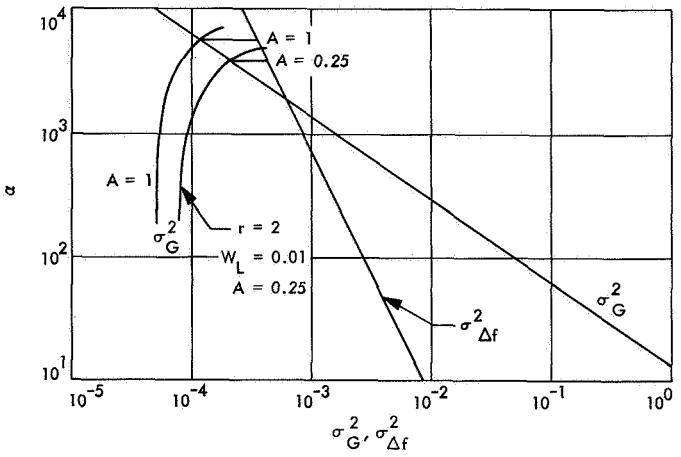


Fig. 7. Graphical solution to nonlinear frequency tracking loop

9. Discrete Equivalent of the Tracking Loop

The actual MFSK receiver updates its estimate of the frequency error only once every T seconds, requiring that the error remain small between these discrete sampling times. Three different discrete equivalent tracking loop filters were tested in a loop simulation program on the digital computer using a Monte Carlo technique. The three loops are shown in Fig. 8. The objective is to obtain an interpolated error function at a rate ten times faster than the calculation of the estimate. The resulting variance of the error Δf for the three different versions is

Version a: $\Delta f = 3.1 \times 10^{-4}$

Version b: $\Delta f = 2.9 \times 10^{-4}$

Version c: $\Delta f = 3.2 \times 10^{-4}$

From these data, it was concluded that version b offered the best solution, and this technique was implemented in the overall MFSK simulation program.

10. Experimental Results

The frequency tracking loop design described in *Subsection 9* has been included in the MFSK simulation program. In addition, an MFSK transmitter simulator has been designed and built for the purpose of removing from the central computer the functions of choosing and setting the transmitter frequency. Modifications were made to the program so that the central computer simulates only the receiver and is capable of operating in any one of the following modes:

(1) Time and frequency sync fixed.

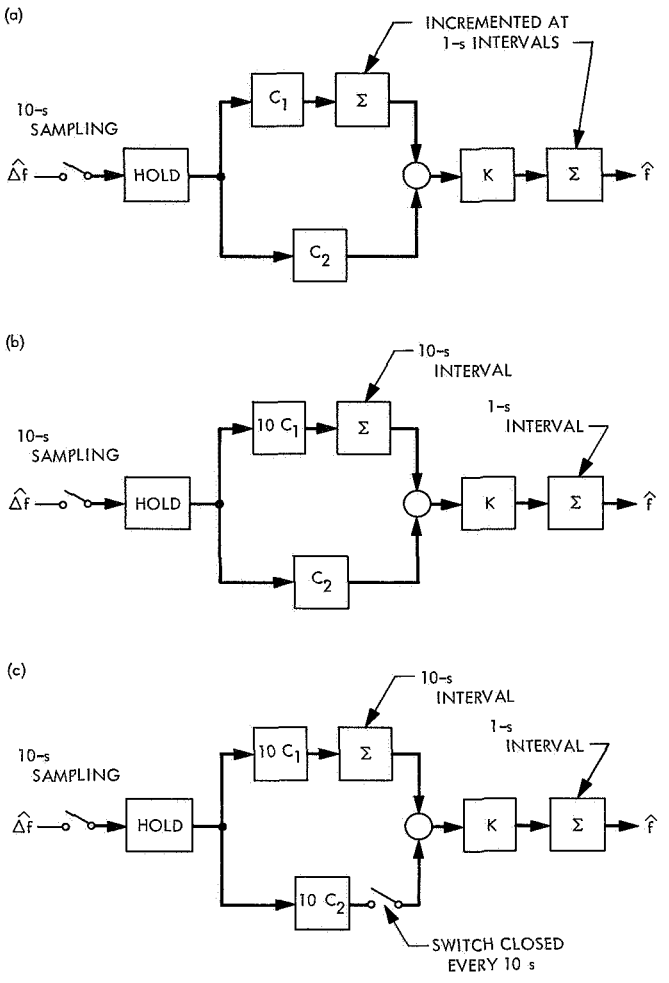


Fig. 8. Three versions of discrete equivalent to filter $H(s)$

(2) Time sync fixed, frequency sync by tracking in the computer.

(3) Frequency sync fixed, time sync by tracking in the computer.

(4) Time and frequency sync tracking in the computer (full operating mode).

For the purpose of testing the frequency tracking loop, mode 2 was used. However, mode 4 has also been tested satisfactorily, showing that the complete MFSK system, with both time and frequency tracking, is feasible. Further testing is required, however, to determine the complete operating characteristics of the system under various possible conditions.

Experimental runs were made both with a fixed input frequency and with a varying frequency. The fixed frequency test resulted in the variance of the frequency error at $E/N_0 = 12$ dB being

$$\sigma^2_{\Delta f} = 0.0004$$

The second test was made using a sinusoidally frequency-modulated input of peak modulation amplitude ± 0.05 Hz, and a modulation frequency of 0.008 Hz. The resulting measured performance loss was 2 dB for the equivalent probability of error ($P_E^W = 0.01$). It should be noted that this input modulation has a maximum slope of 0.0025 Hz/s but is probably more severe than would be encountered under realistic operating conditions, in that it varies at a much higher frequency than would be expected for long-term drift.

II. Conclusions

The experimental results show that an MFSK receiver, using frequency as well as time synchronization tracking, may be implemented to perform at signal-to-noise ratios (E/N_0) at least as low as 14 dB. This article has outlined the design of a frequency tracking loop for a particular set of design parameters. The loop has been shown to perform within specification in an operating MFSK receiver.

Considerably more testing, on a system basis, is necessary to determine the characteristics of the system, as a whole, and to optimize the design parameters of the complete system.

References

1. Ferguson, M., "Communication at Low Data Rate: Spectral Analysis Receivers," *IEEE Trans. Commun. Technol.*, Vol. COM-16, No. 5, Oct. 1968.
2. Lindsey, W., *Nonlinear Analysis and Synthesis of Generalized Tracking Systems*, Technical Report USCEE317. Department of Electrical Engineering, Electronics Sciences Laboratory, University of Southern California, Los Angeles, Calif., Dec. 1968.
3. Viterbi, A., *Principles of Coherent Communications*. McGraw-Hill Book Co., Inc., New York, N.Y., 1966.

B. Interference-to-Signal Ratios in a Frequency-Multiplexed PM Communication System, M. A. Koerner

I. Introduction

In some communication systems, it is often necessary to transmit several channels of data through the same radio communication link. Frequency-multiplexed, phase-modulation (PM) communication systems are often used for this task. Each data signal is phase-modulated or

frequency-modulated on a sinusoidal subcarrier signal and the sum of these subcarrier signals phase-modulates an RF carrier. A phase-locked-loop receiver is used to track the carrier component of the signal received from the transmitter and coherently demodulate the subcarrier signals. Figure 9 is a simplified block diagram of such a system.

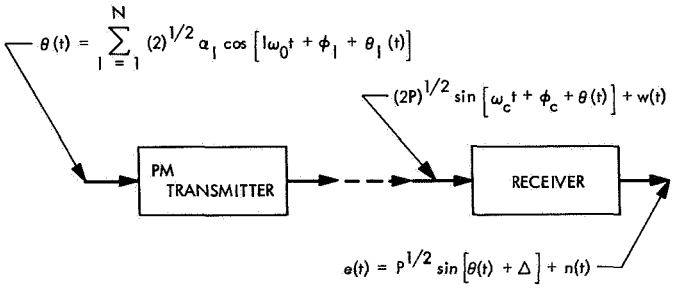


Fig. 9. Frequency-multiplexed PM communication system block diagram

Figure 9 also introduces some of the notation used in this analysis. $\theta(t)$ is the phase modulation carried by the RF signal. P is the received signal power, $f_c = (2\pi)^{-1}\omega_c$ and ϕ_c are the frequency and phase of the RF carrier, and $w(t)$ represents the receiving system noise, a white gaussian process with one-sided power spectral density Φ. $n(t)$ is the noise component of the receiver output. $n(t)$ will have one-sided power spectral density Φ over the frequency band occupied by $P^{1/2}\sin[\theta(t) + \Delta]$. Δ is the receiver phase error. This analysis assumes that the time-varying component of the receiver phase error can be neglected.

As a result of both the net nonlinearity in this modulation/demodulation and the receiver phase error, $P^{1/2}\sin[\theta(t) + \Delta]$ is recovered instead of $\theta(t)$. Hence, numerous intermodulation products between the components of $\theta(t)$ will appear at the receiver output. Provided efficient use of the available RF bandwidth is not a constraint, the subcarrier frequencies can be chosen to achieve any given interference-to-signal ratio in each subcarrier channel. In instances where the option of selecting subcarrier frequencies to achieve acceptable interference-to-signal ratios is not available, the requirement to achieve a given interference-to-signal ratio in each channel may be an important limitation on the efficiency of this type of communication system.

This analysis determines the efficiency and the interference-to-signal ratios which will be obtained in a frequency-multiplexed *PM* communication system carrying N subcarriers. The analysis assumes that the frequency spacing between adjacent subcarriers is $f_0 = (2\pi)^{-1}\omega_0$ and that the first subcarrier has frequency f_0. Hence, the subcarrier frequencies are $f_0, 2f_0, 3f_0, \cdots, Nf_0$ and

$$\theta(t) = \sum_{l=1}^{N} 2^{1/2}\alpha_l \cos[l\omega_0 t + \phi_l + \theta_l(t)] \qquad (1)$$

α_l is the rms phase deviation of the carrier produced by the lth subcarrier, ϕ_l is the initial phase of the lth subcarrier, and $\theta_l(t)$ is the phase modulation carried by the lth subcarrier. This analysis assumes that $\phi_1, \cdots, \phi_N$, and ϕ_c are a set of statistically independent random variables uniformly distributed on $(-\pi, \pi)$.

Since the subcarrier frequencies are harmonically related, the frequencies of the intermodulation products will be equal to a subcarrier frequency or will fall outside the frequency band of interest. While the frequencies of a subcarrier and one or more intermodulation products may be the same, the subcarrier signal will carry only the phase modulation for that channel while each interfering intermodulation product will carry phase modulation from at least one other subcarrier channel. This analysis determines the fraction of the received signal power that is in each subcarrier signal and that which is in all the intermodulation products having the same frequency as the subcarrier signal. Having determined these parameters, an interference-to-signal ratio can be computed for each channel.

2. Subcarrier Signal Power

From Fig. 9, the receiver output is

$$e(t) = P^{1/2}\sin[\theta(t) + \Delta] + n(t) \qquad (2)$$

where $\theta(t)$ is specified by Eq. (1). From analysis reported in SPS 37-39, Vol. IV, pp. 182–192, it can be easily shown that

$$P^{1/2}\sin[\theta(t) + \Delta] = \sum_{n_1=-\infty}^{\infty}, \cdots, \sum_{n_N=-\infty}^{\infty} P^{1/2}\prod_{l=1}^{N} J_{|n_l|}(2^{1/2}\alpha_l)\sin\left(\Delta + \frac{\pi}{2}\sum_{l=1}^{N}|n_l|\right)\cos\left[\left(\sum_{l=1}^{N} l n_l\right)\omega_0 t + \sum_{l=1}^{N} n_l\phi_l + \sum_{l=1}^{N} n_l\theta_l(t)\right] \qquad (3)$$

Noting that terms for which $n_k = \pm 1$, $n_l = 0$, $l \neq k$, $k = 1, 2, \cdots, N$ reproduce the original subcarrier signals, at the receiver output the subcarrier signal in the kth channel will be

$$s_k(t) = P^{1/2} 2J_1(2^{1/2}\alpha_k) \prod_{\substack{l=1 \\ l\neq k}}^{N} J_0(2^{1/2}\alpha_l) \cos(\Delta) \cos[k\omega_0 t + \phi_k + \theta_k(t)] \tag{4}$$

Therefore, the ratio of P_{s_k}, the subcarrier signal power in the kth channel, to P is

$$\frac{P_{s_k}}{P} = 2J_1^2(2^{1/2}\alpha_l) \prod_{\substack{l=1 \\ l\neq k}}^{N} J_0^2(2^{1/2}\alpha_l) \cos^2(\Delta) \tag{5}$$

3. Interference-to-Signal Ratios

This analysis assumes that the effect of the spectral spreading produced by $\theta_1(t), \cdots, \theta_N(t)$ on the interference power in a channel can be neglected. Under this assumption, a term in the summation in Eq. (3) will interfer with the kth subcarrier channel only if

$$\sum_{l=1}^{N} l n_l = \pm k \tag{6}$$

With the exception of the two signal terms, for which $n_k = \pm 1$ and $n_l = 0$ for all $l \neq k$, the interference signal in the kth channel will be the sum of those terms of the summation in Eq. (3) meeting the constaint of Eq. (6). Clearly, every term which satisfies Eq. (6) and for which

$$\sum_{\substack{l=1 \\ l\neq k}}^{N} |n_l| \neq 0 \tag{7}$$

contributes to the interference in the kth channel. The only terms which satisfy Eq. (6) and do not satisfy Eq. (7) are the two signal terms. Hence, the interference signal in the kth channel is

$$i_k(t) = \sum_{\substack{n_1=-\infty \\ \sum_{l=1}^{N} l n_l = \pm k, \\ \sum_{\substack{l=1 \\ l\neq k}}^{N} |n_l| \neq 0}}^{\infty} \cdots \sum_{n_N=-\infty}^{\infty} P^{1/2} \prod_{l=1}^{N} J_{|n_l|}(2^{1/2}\alpha_l) \sin\left(\Delta + \frac{\pi}{2} \sum_{l=1}^{N} |n_l|\right) \cos\left[\left(\sum_{l=1}^{N} l n_l\right)\omega_0 t + \sum_{l=1}^{N} n_l \phi_l + \sum_{l=1}^{N} n_l \theta_l(t)\right] \tag{8}$$

The objective of *Subsections 4* and *5* is to determine P_{i_k}, the total interference power in the kth channel. Having determined P_{s_k} in the preceding subsection, the interference-to-signal P_{i_k}/P_{s_k} in the kth channel can then be computed.

4. Interference Power—Series Solution

The interference power at the receiver output in the kth subcarrier channel is

$$P_{i_k} = \lim_{T\to\infty} (2T)^{-1} \int_{-T}^{T} i_k^2(t)\, dt \tag{9}$$

or, using Eq. (8),

$$P_{i_k} = \sum_{\substack{n_1=-\infty \\ \sum_{l=1}^{N} l n_l = \pm k, \\ \sum_{\substack{l=1 \\ l\neq k}}^{N} |n_l| \neq 0,}}^{\infty} \cdots \sum_{\substack{n_N=-\infty \\ \sum_{l=1}^{N} l m_l = \pm k, \\ \sum_{\substack{l=1 \\ l\neq k}}^{N} |m_l| \neq 0}}^{\infty} \cdots \sum_{m_N=-\infty}^{\infty} P \prod_{l=1}^{N} J_{|n_l|}(2^{1/2}\alpha_l) \prod_{l=1}^{N} J_{|m_l|}(2^{1/2}\alpha_l) \sin\left(\Delta + \frac{\pi}{2} \sum_{l=1}^{N} |n_l|\right) \sin\left(\Delta + \sum_{l=1}^{N} |m_l|\right)$$

$$\times \frac{1}{2} \lim_{T\to\infty} (2T)^{-1} \int_{-T}^{T} \left\{ \cos\left[\sum_{l=1}^{N} l(n_l - m_l)\omega_0 t + \sum_{l=1}^{N} (n_l - m_l)\phi_l + \sum_{l=1}^{N} (n_l - m_l)\theta_l(t)\right] \right.$$

$$\left. + \cos\left[\sum_{l=1}^{N} l(n_l + m_l)\omega_0 t + \sum_{l=1}^{N} (n_l + m_l)\phi_l + \sum_{l=1}^{N} (n_l + m_l)\theta_l(t)\right] \right\} dt \tag{10}$$

In Eq. (10), for every $(m_1, m_2, \cdots, m_N)$ which satisfies the constraints on the summation, $(-m_1, -m_2, \cdots, -m_N)$ will also satisfy the constraints. Hence, the set of values assumed by the set $(n_1 - m_1, n_2 - m_2, \cdots, n_N - m_N)$ is the same as the set of values assumed by $(n_1 + m_1, n_2 + m_2, \cdots, n_N + m_N)$. Thus,

$$P_{i_k} = \sum_{n_1=-\infty}^{\infty}, \cdots, \sum_{n_N=-\infty}^{\infty} \sum_{m_1=-\infty}^{\infty}, \cdots, \sum_{m_N=-\infty}^{\infty} P \prod_{l=1}^{N} J_{|n_l|}(2^{1/2}\alpha_l) \prod_{l=1}^{N} J_{|m_l|}(2^{1/2}\alpha_l) \sin\left(\Delta + \frac{\pi}{2}\sum_{l=1}^{N}|n_l|\right) \sin\left(\Delta + \frac{\pi}{2}\sum_{l=1}^{N}|m_l|\right)$$

$$\sum_{l=1}^{N} l n_l = \pm k, \ \sum_{\substack{l=1 \\ l\neq k}}^{N} |n_l| \neq 0, \ \sum_{l=1}^{N} l m_l = \pm k, \ \sum_{\substack{l=1 \\ l\neq k}}^{N} |m_l| \neq 0$$

$$\times \lim_{T\to\infty} (2T)^{-1} \int_{-T}^{T} \cos\left[\sum_{l=1}^{N} l(n_l - m_l)\omega_0 t + \sum_{l=1}^{N}(n_l - m_l)\phi_l + \sum_{l=1}^{N}(n_l - m_l)\theta_l(t)\right] dt \tag{11}$$

Since $\phi_1, \cdots, \phi_N$ are a set of statistically independent random variables which are uniformly distributed over $(-\pi, \pi)$,

$$\frac{P_{i_k}}{P} = \sum_{n_1=-\infty}^{\infty}, \cdots, \sum_{n_N=-\infty}^{\infty} \prod_{l=1}^{N} J_{|n_l|}^2(2^{1/2}\alpha_l) \sin^2\left(\Delta + \frac{\pi}{2}\sum_{l=1}^{N}|n_l|\right) \tag{12}$$

$$\sum_{l=1}^{N} l n_l = \pm k, \ \sum_{\substack{l=1 \\ l\neq k}}^{N} |n_l| \neq 0$$

While it is theoretically possible to evaluate P_{i_k}/P to any desired accuracy using Eq. (12), this method is rather cumbersome. In the following subsection, a simplified expression is derived for P_{i_k}/P.

5. Interference Power—Integral Solution

Although Eq. (12) would be a numerically inefficient method of evaluating P_{i_k}/P, *Subsection 4* shows that P_{i_k}/P is independent of $\theta_1(t), \cdots, \theta_N(t)$. Hence, for the purpose of computing P_{i_k}/P it may be assumed that $\theta_k(t) = 0$, $k = 1, 2, \cdots, N$. In this case,

$$\theta(t) = \sum_{k=1}^{N} 2^{1/2}\alpha_k \cos(k\omega_0 t + \phi_k) \tag{13}$$

When $\theta(t)$ has the form specified in Eq. (13), the autocorrelation function of $P^{1/2}\sin[\theta(t) + \Delta]$ is

$$\begin{aligned} R_{P^{1/2}\sin[\theta(t)+\Delta]}(\tau) &= \lim_{T\to\infty}(2T)^{-1}\int_{-T}^{T} P\sin[\theta(t) + \Delta]\sin[\theta(t+\tau) + \Delta]\,dt \\ &= \lim_{T\to\infty}(2T)^{-1}\int_{-T}^{T}\frac{1}{2}P\cos[\theta(t+\tau) - \theta(t)]\,dt - \lim_{T\to\infty}(2T)^{-1}\int_{-T}^{T}\frac{1}{2}P\cos[\theta(t+\tau) + \theta(t) + 2\Delta]\,dt \\ &= \frac{1}{2}P\,Re\left(\lim_{T\to\infty}(2T)^{-1}\int_{-T}^{T}\exp\{i[\theta(t+\tau) - \theta(t)]\}\,dt\right) \\ &\quad - \frac{1}{2}P\,Re\left(\exp(i2\Delta)\lim_{T\to\infty}(2T)^{-1}\int_{-T}^{T}\exp\{i[\theta(t+\tau) + \theta(t)]\}\,dt\right) \end{aligned} \tag{14}$$

Using Eq. (13),

$$\begin{aligned} \theta(t+\tau) - \theta(t) &= \sum_{l=1}^{N} 2^{1/2}\alpha_l\{\cos[l\omega_0(t+\tau) + \phi_l] - \cos(l\omega_0 t + \phi_l)\} \\ &= \sum_{l=1}^{N} -2^{3/2}\alpha_l \sin\left(l\omega_0\frac{\tau}{2}\right)\sin\left(l\omega_0 t + l\omega_0\frac{\tau}{2} + \phi_l\right) \end{aligned} \tag{15}$$

and

$$\theta(t+\tau)+\theta(t)=\sum_{l=1}^{N}2^{1/2}\alpha_l\{\cos[l\omega_0(t+\tau)+\phi_l]+\cos(l\omega_0 t+\phi_l)\}$$
$$=\sum_{l=1}^{N}2^{3/2}\alpha_l\cos\left(l\omega_0\frac{\tau}{2}\right)\cos\left(l\omega_0 t+l\omega_0\frac{\tau}{2}+\phi_l\right) \tag{16}$$

Hence,

$$\lim_{T\to\infty}(2T)^{-1}\int_{-T}^{T}\exp\{i[\theta(t+\tau)-\theta(t)]\}\,dt=\lim_{T\to\infty}(2T)^{-1}\int_{-T}^{T}\prod_{l=1}^{N}\exp\left[-i2^{3/2}\alpha_l\sin\left(l\omega_0\frac{\tau}{2}\right)\sin\left(l\omega_0 t+l\omega_0\frac{\tau}{2}+\phi_l\right)\right]dt \tag{17}$$

and

$$\lim_{T\to\infty}(2T)^{-1}\int_{-T}^{T}\exp\{i[\theta(t+\tau)+\theta(t)]\}\,dt=\lim_{T\to\infty}(2T)^{-1}\int_{-T}^{T}\prod_{l=1}^{N}\exp\left[i2^{3/2}\alpha_l\cos\left(l\omega_0\frac{\tau}{2}\right)\cos\left(l\omega_0 t+l\omega_0\frac{\tau}{2}+\phi_l\right)\right]dt \tag{18}$$

Taking the expectation with respect to $\phi_1, \cdots, \phi_N$ of both sides of Eqs. (17) and (18),

$$\lim_{T\to\infty}(2T)^{-1}\int_{-T}^{T}\exp\{i[\theta(t+\tau)-\theta(t)]\}\,dt=\prod_{l=1}^{N}J_0\left[2^{3/2}\alpha_l\sin\left(l\omega_0\frac{\tau}{2}\right)\right] \tag{19}$$

and

$$\lim_{T\to\infty}(2T)^{-1}\int_{-T}^{T}\exp\{i[\theta(t+\tau)+\theta(t)]\}\,dt=\prod_{l=1}^{N}J_0\left[2^{3/2}\alpha_l\cos\left(l\omega_0\frac{\tau}{2}\right)\right] \tag{20}$$

Hence, using Eqs. (19) and (20) in Eq. (14),

$$R_{P^{1/2}\sin[\theta(t)+\Delta]}(\tau)=\frac{1}{2}P\left\{\prod_{l=1}^{N}J_0\left[2^{3/2}\alpha_l\sin\left(l\omega_0\frac{\tau}{2}\right)\right]-\cos(2\Delta)\prod_{l=1}^{N}J_0\left[2^{3/2}\alpha_l\cos\left(l\omega_0\frac{\tau}{2}\right)\right]\right. \tag{21}$$

Expanding $R_{P^{1/2}\sin[\theta(t)+\Delta]}(\tau)$ in a Fourier series,

$$\frac{1}{2}P\left\{\prod_{l=1}^{N}J_0\left[2^{3/2}\alpha_l\sin\left(l\omega_0\frac{\tau}{2}\right)\right]-\cos(2\Delta)\prod_{l=1}^{N}J_0\left[2^{3/2}\alpha_l\cos\left(l\omega_0\frac{\tau}{2}\right)\right]\right\}=\sum_{k=0}^{\infty}P_k\cos(k\omega_0\tau) \tag{22}$$

where

$$P_k=\frac{2}{T_0}\int_{-T_0/2}^{T_0/2}\frac{1}{2}P\left\{\prod_{l=1}^{N}J_0\left[2^{3/2}\alpha_l\sin\left(l\omega_0\frac{\tau}{2}\right)\right]-\cos(2\Delta)\prod_{l=1}^{N}J_0\left[2^{3/2}\alpha_l\cos\left(l\omega_0\frac{\tau}{2}\right)\right]\right\}\cos(k\omega_0\tau)\,d\tau \tag{23}$$

Hence, P_k is the total power—signal plus interference—in the kth channel. Simplifying Eq. (23),

$$\frac{P_k}{P}=\pi^{-1}\int_0^{\pi}\prod_{l=1}^{N}J_0\left[2^{3/2}\alpha_l\sin\left(\frac{l\theta}{2}\right)\right]\cos(k\theta)\,d\theta-\cos(2\Delta)\,\pi^{-1}\int_0^{\pi}\prod_{l=1}^{N}J_0\left[2^{3/2}\alpha_l\cos\left(\frac{l\theta}{2}\right)\right]\cos(k\theta)\,d\theta \tag{24}$$

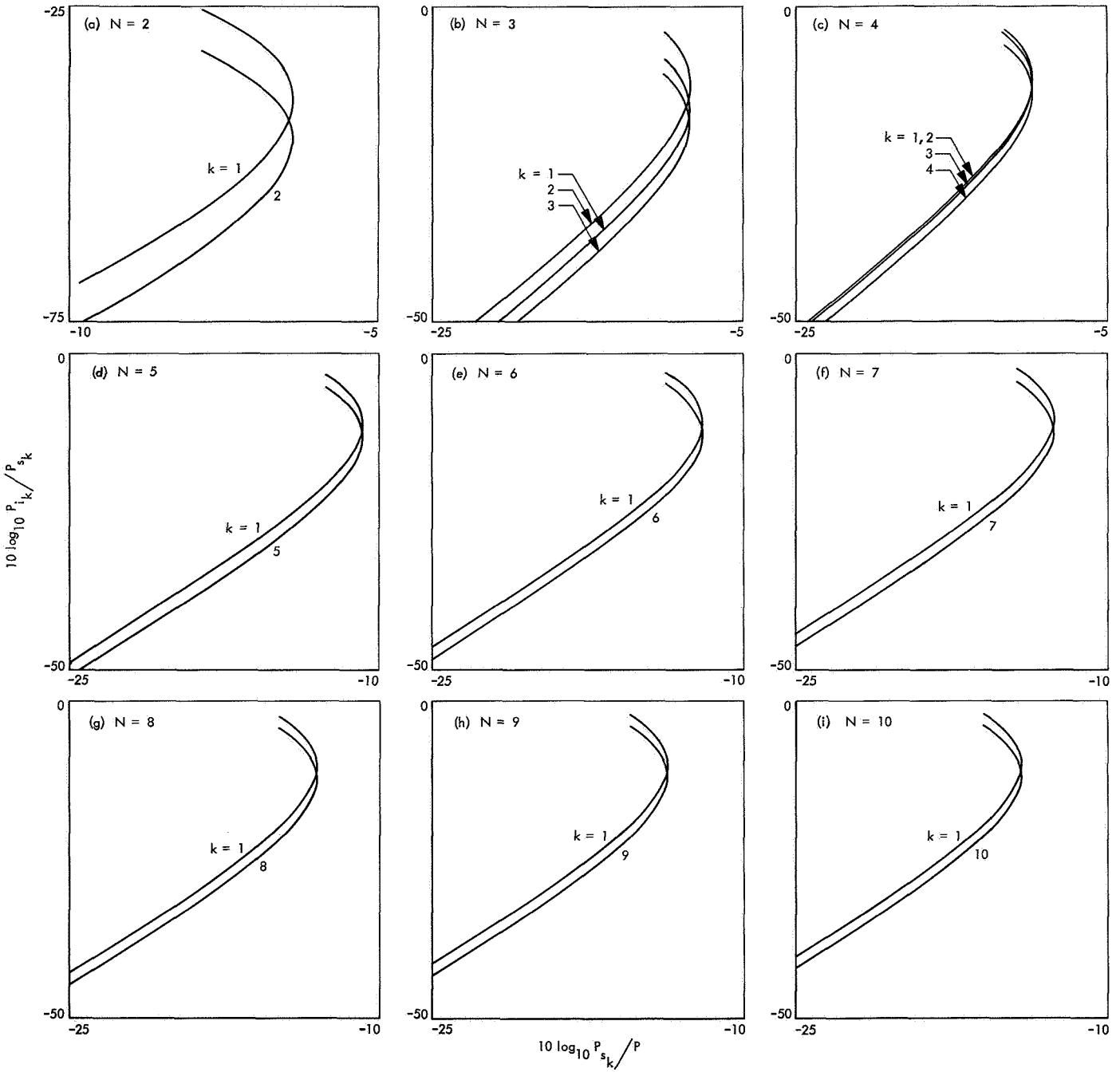


Fig. 10. $10 \log_{10} (P_{i_k}/P_{s_k})$ as a function of $10 \log_{10} (P_{s_k}/P)$ for $N = 2, \cdots, 10$ and $\Delta = 0.0$ deg

Then,

$$\frac{P_{i_k}}{P} = \frac{P_k}{P} - \frac{P_{s_k}}{P} \tag{25}$$

6. Numerical Results

The numerical results presented in this article assume that each of the N subcarriers produces the same rms phase deviation of the carrier. Thus,

$$\alpha_l = \frac{\theta_{\text{rms}}}{N^{1/2}}, \qquad l = 1, 2, \cdots, N \tag{26}$$

where θ_{rms} is the total rms phase deviation of the carrier produced by the N subcarriers. In this case,

$$\frac{P_{s_k}}{P} = 2J_1^2\left[\left(\frac{2}{N}\right)^{1/2}\theta_{\text{rms}}\right]J_0^{2(N-1)}\left[\left(\frac{2}{N}\right)^{1/2}\theta_{\text{rms}}\right] \tag{27}$$

is the same for all channels. The interference power, however, varies from channel to channel, with less interference power in the higher channels than in the lower channels. Using Eq. (26) in Eq. (24),

$$\frac{P_{s_k}}{P} + \frac{P_{i_k}}{P} = \pi^{-1}\int_0^{\pi}\cos(k\theta)\prod_{l=1}^{N}J_0\left[2\left(\frac{2}{N}\right)^{1/2}\theta_{\text{rms}}\sin\left(\frac{l\theta}{2}\right)\right]d\theta - \cos(2\Delta)\,\pi^{-1}\int_0^{\pi}\cos(k\theta)\prod_{k=1}^{N}J_0\left[2\left(\frac{2}{N}\right)^{1/2}\theta_{\text{rms}}\cos\left(\frac{l\theta}{2}\right)\right]d\theta \tag{28}$$

Figures 10a through 10i show the interference-to-signal ratio which can be expected as a function of P_{s_k}/P for a frequency-multiplexed, *PM* communication system which uses the subcarrier frequencies assumed for this analysis. Data is presented for systems having two to ten channels. For systems having two, three, or four channels, 10 $\log_{10}(P_{i_k}/P_{s_k})$ is presented as a function of 10 $\log_{10}(P_{s_k}/P)$ for each channel. For systems having five to ten channels, this data is presented only for the lowest ($k = 1$) and highest ($k = N$) channels. The curves presented in these figures were produced by computing 10 $\log_{10}(P_{i_k}/P_{s_k})$ and 10 $\log_{10}(P_{s_k}/P)$ as a function of θ_{rms} and plotting the results parametrically.

Figure 11 shows an example of the degradation which may be expected when Δ, the receiver phase error, is not zero. For a system having five channels ($N = 5$), the interference-to-signal ratio in the lowest channel ($k = 1$) is plotted as a function of 10 $\log_{10}(P_{s_k}/P)$ for phase errors of 0, 10, 20, and 30 deg.

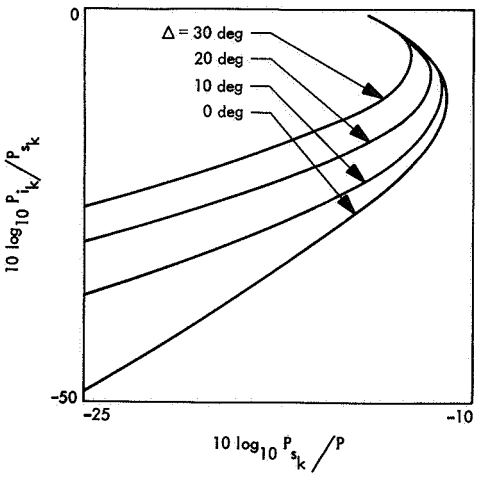


Fig. 11. 10 $\log_{10}(P_{i_k}/P_{s_k})$ as a function of 10 $\log_{10}(P_{s_k}/P)$ for $N = 5$ and $k = 1$ with Δ as the curve parameter

7. Conclusion

The graphs presented in Figs. 10a through 10i indicate that the interference-to-signal ratios which can be obtained for reasonable values of 10 $\log_{10}(P_{s_k}/P)$ in a frequency-multiplexed, *PM* communication system using the subcarrier frequencies specified in this article may be acceptable for some applications. For example, Fig. 10i shows that at least a -20.0 dB interference-to-signal ratio can be obtained in each channel of a ten-channel system while obtaining about 25% of the received signal power in the subcarrier signals. The curves in Figs. 10a to 10i also indicate that reducing the interference-to-signal ratio significantly below -20.0 dB will require operating the multiplexing system at a substantially lower efficiency.

C. Interference-to-Signal Ratios in a PM Single-Sideband Transmitter, *M. A. Koerner*

1. Introduction

It is often necessary to simultaneously transmit several channels of data through a communication system. In

situations where the RF bandwidth must be used as efficiently as possible, the use of double-sideband transmission, such as that discussed in the preceding article, is not acceptable. Furthermore, in communication systems in which RF bandwidth must be used as efficiently as possible, it is usually not feasible to obtain the 1% frequency separation between adjacent channels which is required for conventional multiplexing circuitry.

One system which might be used under these circumstances is shown in Fig. 12. Figure 12 also introduces much of the notation which is used in this analysis. A signal,

$$\theta(t) = \sum_{l=1}^{N} 2^{1/2}\alpha_l \cos\left[l\omega_0 t + \phi_l + \theta_l(t)\right] \qquad (1)$$

phase-modulates a RF carrier which has frequency $f_c = (2\pi)^{-1}\omega_c$ and initial phase ϕ_c. The frequencies of these modulating signals are chosen so that the frequency spacing between adjacent channels is $f_0 = (2\pi)^{-1}\omega_0$ and the first signal has frequency f_0. α_l is the rms phase deviation of the carrier produced by the lth modulating signal. ϕ_l is the initial phase of the lth modulating signal. $\theta_l(t)$ is the net phase modulation carried by the lth modulating signal. This analysis assumes that the initial phases $\phi_1, \cdots, \phi_N$, and ϕ_c are a set of statistically independent random variables uniformly distributed over $(-\pi, \pi)$. The resulting phase-modulated carrier is amplified. P is the total power in the phase-modulated sine wave at the amplifier output. A filter with the amplitude response shown is Fig. 13 removes the carrier component and lower sideband of $e(t)$ as well as any intermodulation products which fall outside the band $[f_c + \frac{1}{2}f_0, f_c + (N + \frac{1}{2})f_0]$ which is used by the N RF channels.

The system shown in Fig. 12 is very similar to a conventional single-sideband transmitter. The principal differences between a single-sideband transmitter and the system shown in Fig. 12 are that (1) phase modulation is used instead of amplitude modulation, and (2) the filter which removes the out-of-band components of $e(t)$ is placed after the power amplifier instead of between the modulator and the power amplifier. The advantage of this mechanization is that the signal to be amplified is a single, phase-modulated sine wave. Such a signal can be amplified in a nonlinear amplifier without further distortion beyond that which is inherent in the phase modulator. Hence, the interference-to-signal ratio and the fraction of the amplifier power output in each RF channel will be independent of the power-amplifier transfer characteristics. Furthermore, for a given raw power input,

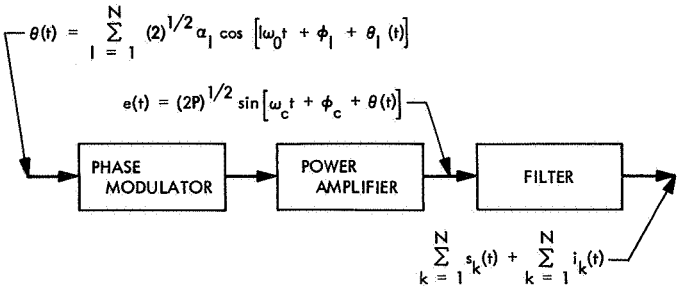


Fig. 12. PM single-sideband transmitter block diagram

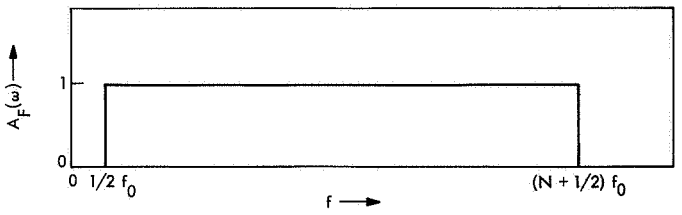


Fig. 13. Output filter amplitude response

the total signal power that can be delivered by a power amplifier such as a traveling-wave tube is substantially greater (6 to 12 dB depending on the crest factor of the input signal) for an amplifier operating near saturation than for an amplifier which must have a linear transfer function. The principal disadvantage of this mechanization is the loss of over 50% of the amplifier power output in the filter which removes the carrier component and lower sideband of $e(t)$.

This analysis determines the efficiency and the interference-to-signal ratios which will be obtained in the N RF channels of the phase-modulation (PM) single-sideband transmitter shown in Fig. 12. The analysis assumes that the frequency spacing between adjacent channels is f_0 and that the first channel is f_0 Hz above the carrier frequency. Hence, the frequencies of the N RF channels are $f_c + f_0, f_c + 2f_0, \cdots, f_c + Nf_0$. It should be noted that this article assumes the upper sideband is used; the interference-to-signal ratios and the fraction of the amplifier power output in the signal in each RF channel would be the same if the lower sideband was used.

Since the frequencies of the modulating signals are harmonically related, the frequency of an intermodulation product at the transmitter output will either be equal to the signal frequency for one of the RF channels or will fall outside the frequency band of interest. While the frequency of the signal in a channel will be the same as that of the intermodulation products in the channel, the signal

will carry only the phase modulation for that channel while each interfering intermodulation product will carry phase modulation from at least one other channel. This analysis determines the fraction of the amplifier power output in the signal in each RF channel and that which is in all the intermodulation products having the same frequency as this signal. Having determined these parameters, interference-to-signal ratios can be computed for each RF channel.

2. Subcarrier Signal Power

From Fig. 12, the power amplifier output is

$$e(t) = (2P)^{1/2} \sin [\omega_c t + \phi_c + \theta(t)] \tag{2}$$

where $\theta(t)$ is specified in Eq. (1). From analysis reported in SPS 37-39, Vol. IV, pp. 182–192, it can be easily shown that

$$e(t) = \sum_{n_1=-\infty}^{\infty}, \cdots, \sum_{n_N=-\infty}^{\infty} (2P)^{1/2} \prod_{l=1}^{N} J_{n_l}(2^{1/2}\alpha_l) \sin\left[\left(\omega_c + \sum_{l=1}^{N} l n_l \omega_0\right) t + \phi_c + \sum_{l=1}^{N} n_l \phi_l + \sum_{l=1}^{N} n_l \theta_l(t)\right] \tag{3}$$

The terms for which $n_k = 1, n_l = 0, l \neq k, k = 1, 2, \cdots, N$ reproduce the original modulating signals shifted up in frequency by f_c. The signal in the kth channel will be

$$s_k(t) = (2P)^{1/2} J_1(2^{1/2}\alpha_k) \prod_{\substack{l=1 \\ l\neq k}}^{N} J_0(2^{1/2}\alpha_l) \sin[(\omega_c + k\omega_0) t + \phi_c + \phi_k + \theta_k(t)] \tag{4}$$

Therefore, the ratio of P_{s_k}, the signal power in the kth channel, to P is

$$\frac{P_{s_k}}{P} = J_1^2(2^{1/2}\alpha_k) \prod_{\substack{l=1 \\ l\neq k}}^{N} J_0^2(2^{1/2}\alpha_l) \tag{5}$$

3. Interference-to-Signal Ratios

This analysis assumes that the effect of the spectral spreading produced by $\theta_1(t), \cdots, \theta_N(t)$ on the interference power in a channel can be neglected. Under this assumption, a term in the summation in Eq. (3) will interfere with the kth RF channel only if

$$\sum_{l=1}^{N} l n_l = k \tag{6}$$

With the exception of the signal term, for which $n_k = 1$, $n_l = 0$, $l \neq k$, the interference signal in the kth RF channel will be the sum of those terms of the summation in Eq. (3) meeting the constraint of Eq. (6). Clearly, every term which satisfies Eq. (6) and for which

$$\sum_{\substack{l=1 \\ l\neq k}}^{N} |n_l| \neq 0 \tag{7}$$

contributes to the interference in the kth channel. The only term which satisfies Eq. (6) and does not satisfy Eq. (7) is the signal term. Hence, the inteference signal in the kth RF channel is

$$i_k(t) = \sum_{\substack{n_1=-\infty \\ \sum_{l=1}^{N} l n_l = k,}}^{\infty}, \cdots, \sum_{\substack{n_N=-\infty \\ \sum_{\substack{l=1 \\ l\neq k}}^{N} |n_l| \neq 0}}^{\infty} (2P)^{1/2} \prod_{l=1}^{N} J_{n_l}(2^{1/2}\alpha_l) \sin\left[(\omega_c + k\omega_0) t + \phi_c + \sum_{l=1}^{N} n_l \phi_l + \sum_{l=1}^{N} n_l \theta_l(t)\right] \tag{8}$$

The objective of *Subsections 4* and *5* is to determine P_{i_k}, the total interference power in the kth channel. Having determined P_{s_k} in the preceding subsection, the interference-to-signal ratio P_{i_k}/P_{s_k} in the kth channel can then be computed.

4. Interference Power—Series Solution

The interference power at the power amplifier output in the kth RF channel is

$$P_{i_k} = \lim_{T\to\infty} (2T)^{-1} \int_{-T}^{T} i_k^2(t)\, dt \tag{9}$$

or using Eq. (7),

$$\begin{aligned} P_{i_k} = {} & \sum_{n_1=-\infty}^{\infty}, \cdots, \sum_{n_N=-\infty}^{\infty} \sum_{m_1=-\infty}^{\infty}, \cdots, \sum_{m_N=-\infty}^{\infty} P \prod_{l=1}^{N} J_{n_l}(2^{1/2}\alpha_l) \prod_{l=1}^{N} J_{m_l}(2^{1/2}\alpha_l) \lim_{T\to\infty} (2T)^{-1} \\ & \sum_{l=1}^{N} l n_l = k, \sum_{\substack{l=1 \\ l\neq k}}^{N} |n_l| \neq 0, \quad \sum_{l=1}^{N} l m_l = k, \sum_{\substack{l=1 \\ l\neq k}}^{N} |m_l| \neq 0 \\ & \times \int_{-T}^{T} \left\{ \cos\left[\sum_{l=1}^{N} (n_l - m_l)\phi_l + \sum_{l=1}^{N} (n_l - m_l)\theta_l(t) \right] \right. \\ & \left. - \cos\left[2(\omega_c + k\omega_0)t + 2\phi_c + \sum_{l=1}^{N} (n_l + m_l)\phi_l + \sum_{l=1}^{N} (n_l + m_l)\theta_l(t) \right] \right\} dt \end{aligned} \tag{10}$$

While this analysis assumes that the frequencies of the N sine waves in the modulation $\theta(t)$ are harmonically related, it also assumes that $\phi_1, \phi_2, \cdots, \phi_N$, and ϕ_c are statistically independent and uniformly distributed on $(-\pi, \pi)$. Hence, the fraction of the amplifier power output in the kth channel interference is

$$\frac{P_{i_k}}{P} = \sum_{n_1=-\infty}^{\infty}, \cdots, \sum_{n_N=-\infty}^{\infty} \prod_{l=1}^{N} J_{n_l}^2(2^{1/2}\alpha_l) \tag{11}$$
$$\sum_{l=1}^{N} l n_l = k, \sum_{\substack{l=1 \\ l\neq k}}^{N} |n_l| \neq 0$$

While it is theoretically possible to evaluate P_{i_k}/P to any desired accuracy using Eq. (11), this method is rather cumbersome. In the following subsection, a simplified expression is derived for P_{i_k}/P.

5. Interference Power—Integral Solution

Although Eq. (11) would be a numerically inefficient method for evaluating P_{i_k}/P, *Subsection 4* shows that P_{i_k}/P is independent of $\theta_k(t), k = 1, \cdots, N$. Hence, for the purposes of computing P_{i_k}/P it may be assumed that $\theta_k(t) = 0$, $k = 1, 2, \cdots, N$. In this case,

$$\theta(t) = \sum_{k=1}^{N} 2^{1/2}\alpha_k \cos(k\omega_0 t + \phi_k) \tag{12}$$

is a periodic waveform.

When $\theta(t)$ has the form specified in Eq. (12), the autocorrelation function of the power amplifier output is, using Eq. (2),

$$\begin{aligned} R_l(\tau) &= \lim_{T\to\infty} (2T)^{-1} \int_{-T}^{T} (2P)^{1/2} \sin[\omega_c t + \phi_c + \theta(t)]\, (2P)^{1/2} \sin[\omega_c(t+\tau) + \phi_c + \theta(t+\tau)]\, dt \\ &= \lim_{T\to\infty} (2T)^{-1} \int_{-T}^{T} P \cos[\omega_c \tau + \theta(t+\tau) - \theta(t)]\, dt \\ &= P\, Re\left(\exp(i\omega_c \tau) \lim_{T\to\infty} (2T)^{-1} \int_{-T}^{T} \exp\{i[\theta(t+\tau) - \theta(t)]\}\, dt \right) \end{aligned} \tag{13}$$

However, from Eq. (12),

$$\theta(t+\tau)-\theta(t)=\sum_{l=1}^{N}2^{1/2}\alpha_l\{\cos[l\omega_0(t+\tau)+\phi_l]-\cos(l\omega_0 t+\phi_l)\}$$

$$=\sum_{l=1}^{N}-2^{3/2}\alpha_l\sin\left(l\omega_0\frac{\tau}{2}\right)\sin\left(l\omega_0 t+l\omega_0\frac{\tau}{2}+\phi_l\right) \tag{14}$$

Hence,

$$\lim_{T\to\infty}(2T)^{-1}\int_{-T}^{T}\exp\{i[\theta(t+\tau)-\theta(t)]\}\,dt=\lim_{T\to\infty}(2T)^{-1}\int_{-T}^{T}\prod_{l=1}^{N}\exp\left[-i2^{3/2}\alpha_l\sin\left(l\omega_0\frac{\tau}{2}\right)\sin\left(l\omega_0 t+l\omega_0\frac{\tau}{2}+\phi_l\right)\right]dt \tag{15}$$

or, taking the expectation of both sides with respect to $\phi_1, \cdots, \phi_N$,

$$\lim_{T\to\infty}(2T)^{-1}\int_{-T}^{T}\exp\{i[\theta(t+\tau)-\theta(t)]\}\,dt=\lim_{T\to\infty}(2T)^{-1}\int_{-T}^{T}\prod_{l=1}^{N}(2\pi)^{-1}\int_{-\pi}^{\pi}\exp\left[-i2^{3/2}\alpha_l\sin\left(l\omega_0\frac{\tau}{2}\right)\right.$$

$$\left.\times\sin\left(l\omega_0 t+l\omega_0\frac{\tau}{2}+\phi_l\right)\right]d\phi_l\,dt=\prod_{l=1}^{N}J_0\left[2^{3/2}\alpha_l\sin\left(l\omega_0\frac{\tau}{2}\right)\right] \tag{16}$$

Hence, using Eq. (16) in Eq. (13),

$$R_e(\tau)=P\cos(\omega_c\tau)\prod_{l=1}^{N}J_0\left[2^{3/2}\alpha_l\sin\left(l\omega_0\frac{\tau}{2}\right)\right] \tag{17}$$

Expanding

$$P\prod_{l=1}^{N}J_0\left[2^{3/2}\alpha_l\sin\left(l\omega_0\frac{\tau}{2}\right)\right]$$

in a Fourier series,

$$P\prod_{l=1}^{N}J_0\left[2^{3/2}\alpha_l\sin\left(l\omega_0\frac{\tau}{2}\right)\right]=\sum_{k=-\infty}^{\infty}P_k\exp(ik\omega_0\tau) \tag{18}$$

and

$$R_e(\tau)=Re\left\{P\exp(i\omega_c\tau)\prod_{l=1}^{N}J_0\left[2^{3/2}\alpha_l\sin\left(l\omega_0\frac{\tau}{2}\right)\right]\right\}$$

$$=\sum_{k=-\infty}^{\infty}P_k\cos[(\omega_c+k\omega_0)t] \tag{19}$$

where

$$P_k=PT_0^{-1}\int_{-T_0/2}^{T_0/2}\left\{\prod_{l=1}^{N}J_0\left[2^{3/2}\alpha_l\sin\left(l\omega_0\frac{\tau}{2}\right)\right]\right\}\exp(-ik\omega_0\tau)\,d\tau \tag{20}$$

Hence, P_k is the total power—signal plus interference—in the kth channel. Simplifying Eq. (20),

$$\frac{P_k}{P} = \pi^{-1} \int_0^{\pi} \prod_{l=1}^{N} J_0\left[2^{3/2}\alpha_l \sin\left(\frac{l\theta}{2}\right)\right] \cos(k\theta)\, d\theta \tag{21}$$

Then

$$\frac{P_{i_k}}{P} = \frac{P_k}{P} - \frac{P_{s_k}}{P} \tag{22}$$

6. Numerical Results

The numerical results presented in this article assume that each of the N sinusoidal modulating signals produces the same rms phase deviation of the carrier. Thus,

$$\alpha_l = \frac{\theta_{\text{rms}}}{N^{1/2}}, \qquad l = 1, 2, \cdots, N \tag{23}$$

where θ_{rms} is the total rms phase deviation of the carrier produced by the N modulating signals. In this case,

$$\frac{P_{s_k}}{P} = J_1^2\left[\left(\frac{2}{N}\right)^{1/2}\theta_{\text{rms}}\right] J_0^{2(N-1)}\left[\left(\frac{2}{N}\right)^{1/2}\theta_{\text{rms}}\right] \tag{24}$$

is the same for all channels. The interference power, however, varies from channel to channel, with less interference in the higher channels than the lower channels. Using Eq. (23) in Eq. (21),

$$\frac{P_{s_k}}{P} + \frac{P_{i_k}}{P} = \pi^{-1} \int_0^{\pi} \prod_{l=1}^{N} J_0 \times\left[2\left(\frac{2}{N}\right)^{1/2}\theta_{\text{rms}} \sin\left(\frac{l\theta}{2}\right)\right] \cos(k\theta)\, d\theta \tag{25}$$

Figures 14a through 14i show the interference-to-signal ratio which can be expected as a function of P_{s_k}/P for a frequency-multiplexed PM communication system which uses the modulating signal frequencies assumed for this analysis. Data is presented for systems having two to ten channels. For systems having two to six channels, $10\log_{10}(P_{i_k}/P_{s_k})$ is plotted as a function of $10\log_{10}(P_{s_k}/P)$ for each channel. For systems having seven to ten channels, this data is presented only for the lowest ($k = 1$) and highest ($k = N$) channels. The curves presented in these figures were produced by computing $10\log_{10}(P_{i_k}/P_{s_k})$ and $\log 10_{10}(P_{s_k}/P)$ as a function of θ_{rms} and plotting the results parametrically.

7. Conclusion

The graphs presented in Figs. 14a through 14i indicate that in a PM single-sideband transmitter, using the modulating signal frequencies specified in this article, the interference-to-signal ratios which can be obtained for values of $10\log_{10}(P_{s_k}/P)$ which are greater than -1.0 below its maximum will be greater than -10.0 db, and hence will probably be unacceptable for most applications. In a subsequent article, we shall show that this difficulty can be overcome if the frequencies of the modulating signals are substantially greater than their frequency spacing.

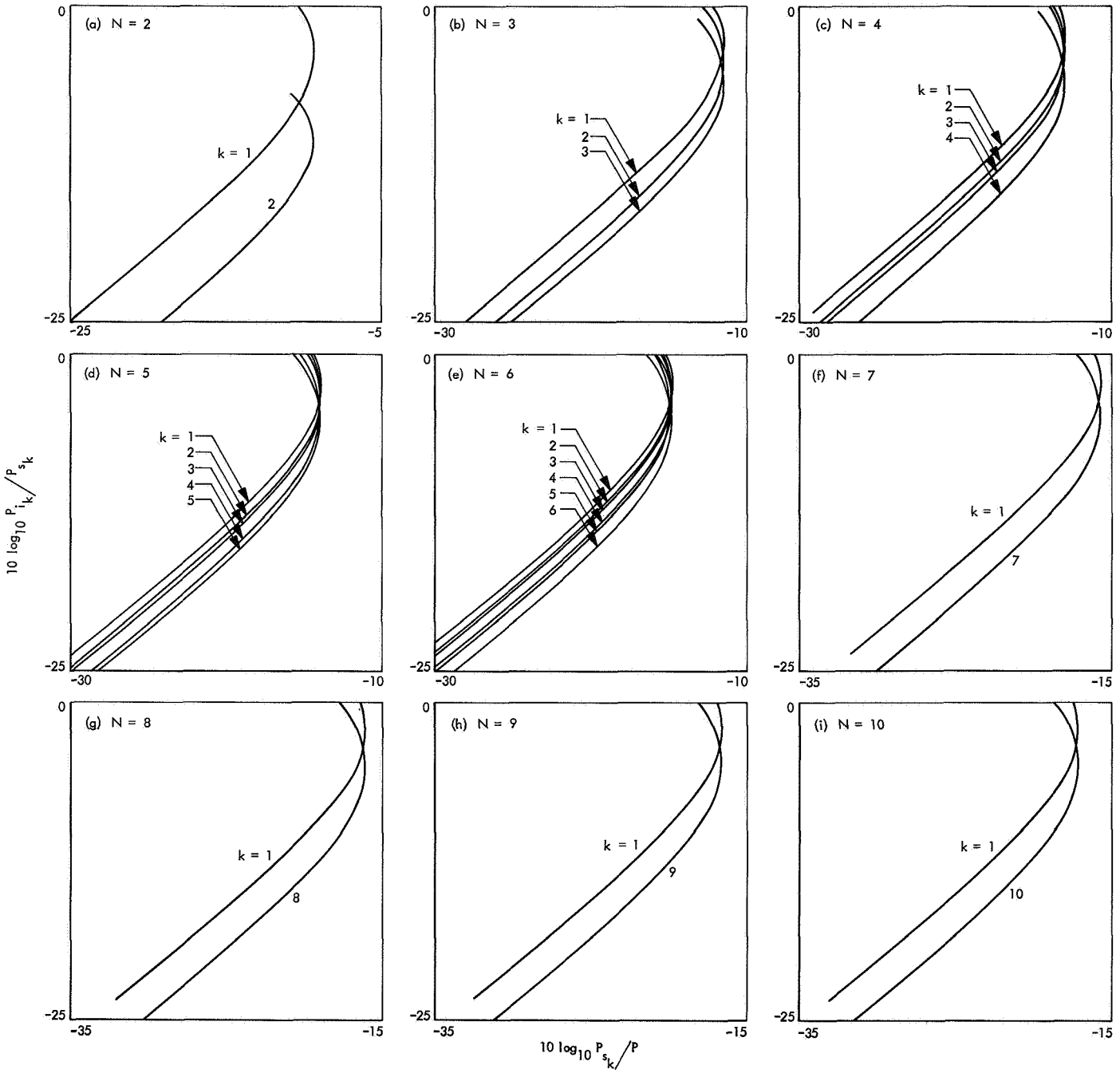


Fig. 14. $10 \log_{10} (P_{i_k}/P_{s_k})$ as a function of $10 \log_{10} (P_{s_k}/P)$ for $N = 2, \cdots, 10$

IX. Spacecraft Power

GUIDANCE AND CONTROL DIVISION

A. Electric Propulsion Power Conditioning,

E. N. Costogue

1. Objective

The electric propulsion power conditioning activities are directed to accomplish two tasks. The first task, SEPST II,[1] has been operational since March 1969 and involves the evaluation testing of a power conditioner supply with two oxide cathode, ion engines. A switching module is utilized to switch power to other engines by command. The second task, SEPST III, is to be operational by late 1969. This task will involve the design, fabrication, and evaluation testing of the complete power conditioning subsystem of the electric propulsion system. The power conditioning subsystem will consist of: (1) one breadboard and two experimental power conditioning units of improved designs with an expected average efficiency of 91–93% and specific weight of 8.6 lb/kW at the operating power level, (2) a switching unit that will be utilized to switch power conditioning units to available ion engines, (3) a maximum power-point monitor to detect the maximum power point of the solar panel source and determine the power margins available to the subsystem, and (4) a solar array simulator to simulate the solar panel output of 15 kW and the voltage–current characteristics for 1 to 5 AU. Items (1) and (4) will be developed under contracts. Items (2) and (3) will be designed and fabricated in-house.

[1]SEPST = solar electric propulsion system test.

2. SEPST II

The SEPST II power conditioner, a modified unit built in 1966 for JPL and later used on the SERT II program,[2] was delivered by the contractor in December 1968. From December until March, integration tests with the ion thruster were conducted and minor modifications had to be introduced to improve the performance of the conditioner and to match its characteristics to the engine demands. The modifications introduced were in the following areas (see Ref. 1 for details):

(1) Overload circuits: The response and tripping levels of the high-voltage supplies were adjusted to what was considered tolerable limits for the engine operation and power conditioner protection.

(2) Thruster closed-loop operational amplifiers: The characteristics of the arc-cathode control-loop operational amplifier were adjusted to the desirable thruster operational levels. The stability of the loop was improved and gain was reduced to an acceptable level to achieve satisfactory operation over the total range of thrust.

(3) Soft recovery from arcing: The recovery after a temporary shutdown of the thruster to interrupt a sustained arc was improved by reducing the magnet output current to a minimum value of 75 mA for 1.5 s and then ramping the current to

[2]SERT = space electric rocket test.

its full value. This modification produced a desirable way to restart the operation of the thrusters.

The switching unit was designed in-house and was delivered on schedule. Some filtering was introduced in the input lines of the unit to eliminate the interference of system-generated noise. The switching unit continues to perform satisfactorily in all the tests conducted with the power conditioner and ground power supplies.

The SEPST II program has completed a 300-h preendurance test. Throughout the test, the power conditioner performance was considered satisfactory, although a number of weak areas in the logic were uncovered. Modifications to the logic and additional circuitry to improve the soft recovery characteristics of the power conditioner after arcing have been introduced. The power conditioner is being prepared for the endurance test.

3. SEPST III

In the SEPST III program, the functional specification for the power conditioning subsystem has been issued.

The activities at Hughes Aircraft Co., the contractor to design, develop, fabricate, and test the power conditioner units, are progressing satisfactorily. Delivery of the breadboard unit is expected by the end of May. The final design review will be held during the third week in June to review the released documents and test results of the unit. A problem in meeting the design goal efficiency of 93% is attributed to poor switching characteristics of the power transistors. Transistors with improved characteristics are being evaluated for possible substitution.

Reference

1. Macie, T., et al., "Solar Electric Propulsion System Evaluation," Paper 69–498, presented at the AIAA 5th Propulsion Joint Specialist Conference, June 9–13, 1969.

B. Required Gamma Ray Shielding for Science Instruments on Radioisotope-Powered Spacecraft, *M. A. Dore*

1. Introduction

The overall objective of the nuclear definition studies is to both analytically and experimentally investigate the problems of integrating a radioisotope thermoelectric generator (RTG) as a source of power to radiation-sensitive spacecraft subsystems. One task presently being undertaken is the analytical determination of required gamma ray shielding for various radiation detectors of the type included in the science experiment package. This is done for three possible spacecraft designs being considered by the thermoelectric outer planet spacecraft (TOPS) project.

Previous work within the Nuclear Power Sources Group at JPL, using the ANISN and QAD-B computer codes, has led to time-dependent 20-group spectra and complete three-dimensional maps of the gamma radiation field for several variations of the basic RTG PuO_2 fuel capsule design anticipated for use with the TOPS spacecraft. Tentative values have been set for (1) the allowable spurious count rate for each of the detectors considered, (2) the separation distance and relative orientation of RTGs to science in each spacecraft configuration considered, (3) the sensitive area for each instrument, and (4) the shield areas needed to effectively shadow the sensitive area of each detector. These values are tentative pending a final, detailed spacecraft design, at which time this analysis will be repeated if necessary.

This information, along with attenuation coefficients for tungsten shields, serves as input data to a special-purpose computer program which generates the required shield thicknesses and weights. Flat, rectangular tungsten shadow shields located near the sensitive instruments are assumed in all cases, and a single shield for the RTGs themselves is considered for one of the configurations.

2. Spacecraft Geometry

Three possible configurations for the TOPS spacecraft were considered for this study (Fig. 1). In configuration 12C, three RTGs are stacked in tandem, and lie directly opposite the science boom. In this way, self-shielding by the plutonium in the RTGs themselves is maximized. In configuration 9B, the three RTGs form a semicircle with the science boom opposite the center one. The relative orientation of RTGs to science here leads to considerably higher fluxes, since self-shielding is not fully utilized nor is the exposed area of the RTG seen by the science package minimized. Configuration 10 assumes only two RTGs, which lie very close to the science package in a highly unfavorable orientation as far as radiation flux is concerned. Note, also, that only configuration 12C offers a small enough solid angle to the science package to consider shielding the RTGs

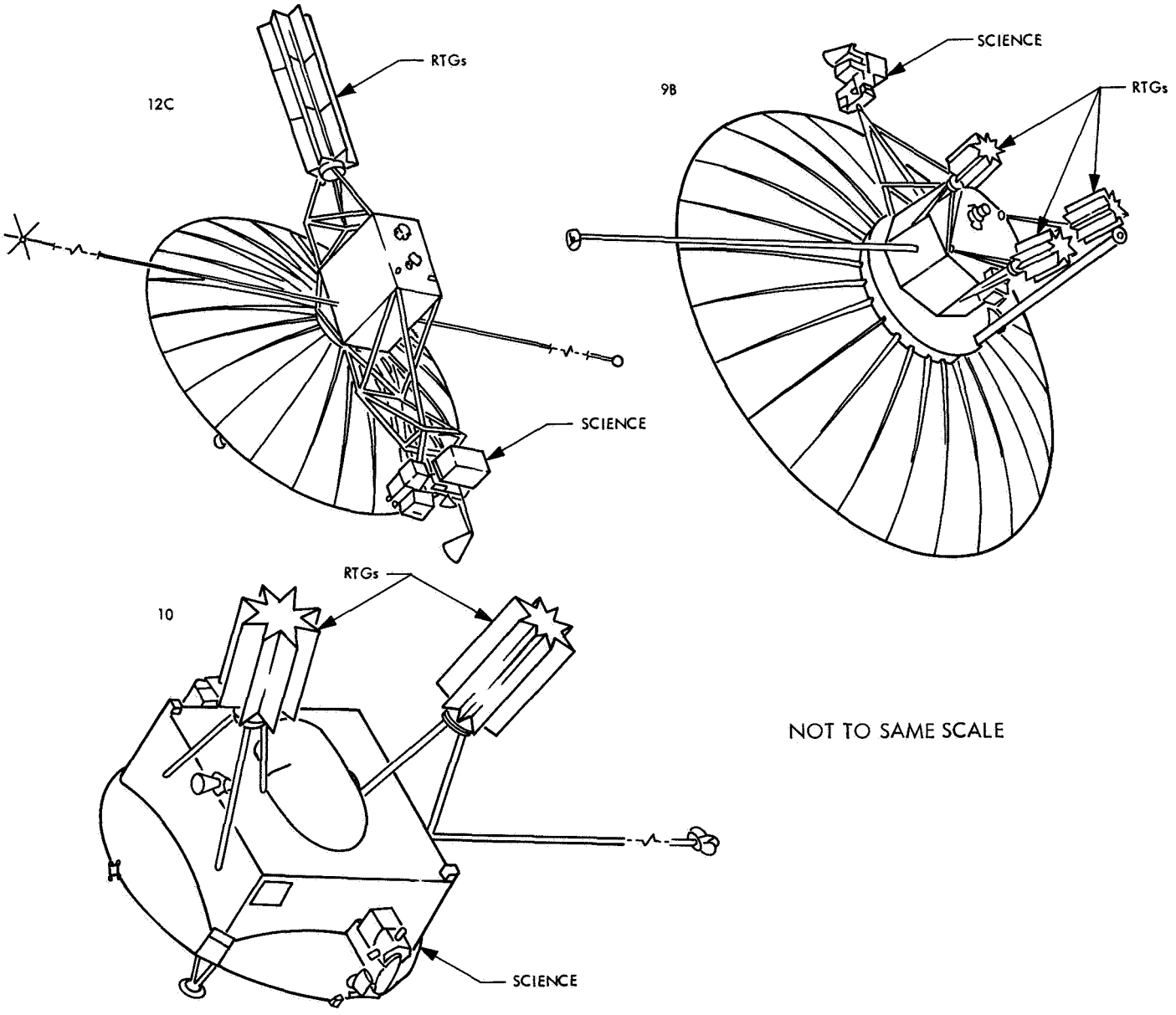


Fig. 1. TOPS configurations

themselves. The most significant geometrical parameters are:

Configuration	Number of RTGs	Power/RTG, W(th)	Separation distance, cm
12C	3	3150	449
9B	3	3150	508 (center) 421 (ends)
10	2	4725	140

In an actual RTG, the fuel capsule is surrounded by the generator, which is made up of high-density thermoelectric materials. Attenuation due to this generator has *not* been considered here, since little reliable information exists on what the actual generator will consist of, or of its detailed geometry. This leads to a worst-case situation, assuming no shielding whatsoever from the generator. Similarly, attenuation provided by the electronics bus is also not considered. After adoption of a definite spacecraft design, both of these factors will be included in a final analysis.

3. Detectors

Four detectors of the general type flown on *Mariner* missions were considered as the critical instruments. They are the Geiger–Müller (G–M) tube EON-6213, the Channeltron electron multiplier CEM-4010, and two surface barrier solid-state detectors, one 31 μm thick, the other 200 μm thick. In all of these, the sensitive area was assumed to be the aperture area plus a $\frac{1}{10}$-in.-wide band around the aperture in which gamma-induced electrons could be formed and still manage to reach sensitive portions of the detector (Fig. 2). The detectors are assumed oriented perpendicular to the incident radiation, and, therefore, represent worst-case situations.

These four detectors were then assumed to be assembled into four science-experiment packages as follows:

LEPEDEA experiment[3]	1 channeltron and 1 G–M tube
Ultraviolet photometer	1 channeltron and 1 G–M tube
Trapped radiation detector	3 G–M tubes and 1 solid-state detector (31 μm)
Cosmic ray telescope	1 solid-state detector (200 μm)

[3]LEPEDEA = low-energy proton and electron differential analyzer.

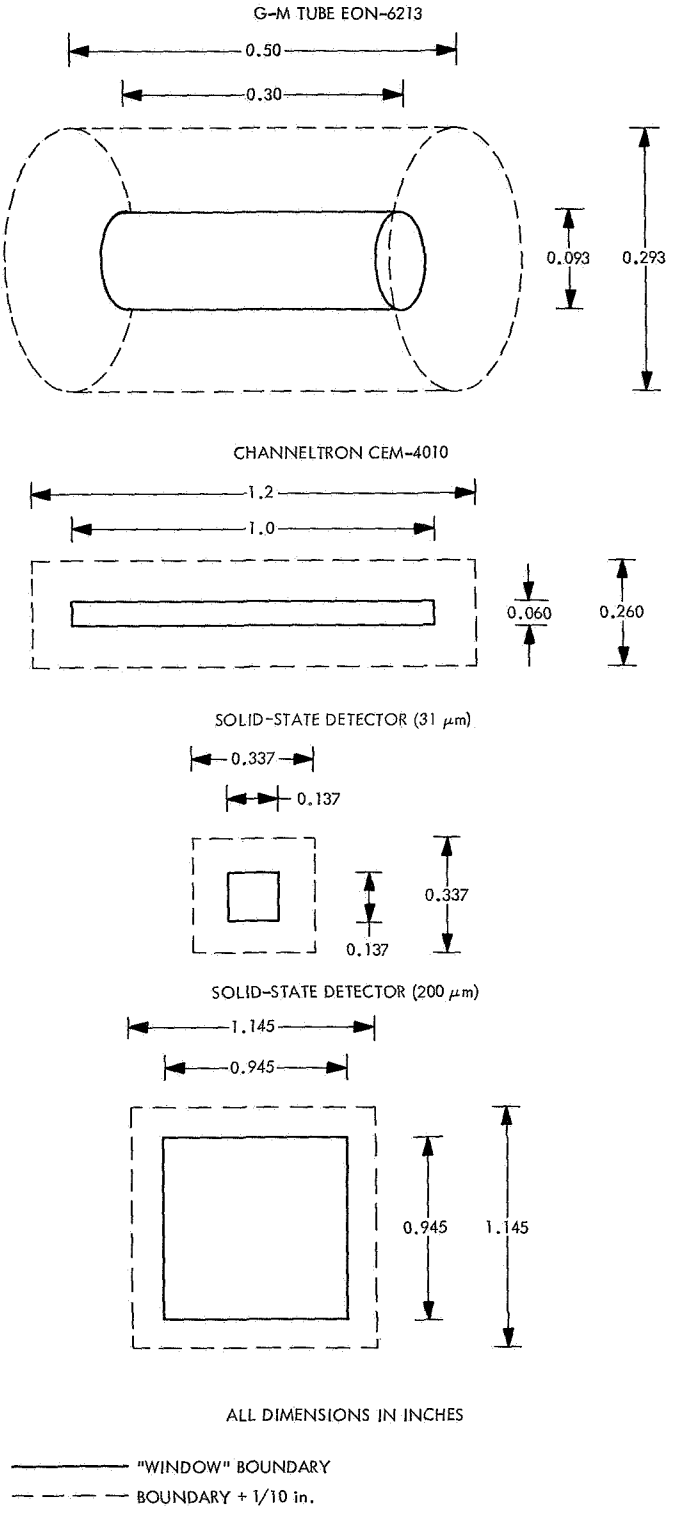


Fig. 2. Assumed sensitive areas of detectors

Detector sensitivities were based upon work done at JPL. Note that the response of the channeltron is assumed identical to that of the G–M tube. This was chosen again as a worst case, since values in the literature give anywhere from 10 to 50% of the G–M tube response for the channeltron.

4. Method

The spectra due to the fuel capsule are in tabulated form for various combinations of radial distance and angular orientation. These are input into the program, which then interpolates to determine the spectra at the position occupied by the science package. Since this is an extended object, the point of highest gamma flux is chosen. This is done for several fuel ages and each spacecraft configuration.

The tungsten shield attenuation coefficients used are taken as constant, there being no allowance made for downscatter in energy since the coefficients were calculated for spectra which are very similar to those employed here and which already include downscatter effects. Also, neutron-induced gammas produced within the shields are not considered. Previous work with PuO_2 neutron and gamma spectra indicates that the contribution of neutron-induced gammas is several orders of magnitude lower than the primary gamma flux which penetrates the shield.

Table 1. Results of experiment–detector analysis (all values are for 18-yr-old fuel)

Parameter	LEPEDEA experiment[a]		Ultraviolet photometer		Trapped radiation detector		Cosmic ray telescope
	Channeltron CEM-4010	G-M tube EON-6213	Channeltron CEM-4010	G-M tube EON-6213	Solid-state detector (12 mm² × 31 μm)	G–M tube EON-6213	Solid-state detector (5.7 cm² × 200 μm)
Sensitive area, cm²	2.01	0.95	2.01	0.95	0.733	0.95	8.48
Maximum allowable spurious rate, counts/s	0.1	0.3	0.1	0.3	0.001	0.3	1.0
Flat shield area, cm	6.0	4.0	10.0	4.0	3.0	2.0 Each	10.0
Shield thickness, in.							
Configuration 12C	2.17	1.32	2.17	1.32	1.45	1.32	0.33
Configuration 9B	2.77	1.92	2.77	1.92	2.31	1.92	1.25
Configuration 10	3.68	2.76	3.68	2.76	3.13	2.76	1.98
Flat shield weight, g							
Configuration 12C	562	228	936	228	188	341	142
Configuration 9B	719	331	1198	331	299	497	541
Configuration 10	953	477	1589	477	406	716	855
Shield weight/experiment, g							
Configuration 12C	789		1164		529		142
Configuration 9B	1050		1529		796		541
Configuration 10	1430		2066		1121		855
Total shield weight, g(lb)							
Configuration 12C	2624 (5.79)						
Configuration 9B	3917 (8.64)						
Configuration 10	5473(12.07)						

[a]LEPEDEA = low-energy proton and electron differential analyzer.

Table 2. Spurious counts/ cm^2-s vs shield thickness for each detector (18-yr-old fuel)

Shield thickness, in.	G–M tube EON-6213	Channeltron CEM-4010	Solid-state detector (31 μm)	Solid-state detector (200 μm)
Configuration 12C				
0.00	5.88	12.48	0.0310	2.30
0.25	3.24	6.87	0.0157	1.17
0.50	1.82	3.87	0.00843	0.627
0.75	1.04	2.20	0.00465	0.346
1.00	0.591	1.25	0.00262	0.195
1.25	0.339	0.719	0.00150	0.112
1.50	0.195	0.414	0.00087	0.0646
1.75	0.113	0.239	0.00051	0.0377
2.00	0.0654	0.139	0.00030	0.0222
2.25	0.0381	0.0808	0.00018	0.0131
2.50	0.0223	0.0473	0.00010	0.00781
2.75	0.0131	0.0278	0.00006	0.00468
Configuration 9B				
0.00	103.03	218.67	10.48	780.2
0.25	13.90	29.49	0.251	18.66
0.50	7.41	15.73	0.115	8.53
0.75	4.05	8.60	0.0543	4.04
1.00	2.26	4.79	0.0267	1.98
1.25	1.27	2.70	0.0135	1.00
1.50	0.727	1.54	0.00701	0.522
1.75	0.418	0.8874	0.00373	0.277
2.00	0.242	0.5142	0.00202	0.150
2.25	0.141	0.300	0.00112	0.0830
2.50	0.0829	0.176	0.00063	0.0465
2.75	0.0490	0.104	0.00036	0.0265
Configuration 10				
0.00	642.2	1363.	65.31	4863.
0.25	86.62	183.8	1.56	116.3
0.50	46.19	98.03	0.714	53.16
0.75	25.26	53.60	0.339	25.21
1.00	14.07	29.85	0.166	12.37
1.25	7.94	16.85	0.0841	6.26
1.50	4.53	9.61	0.0437	3.25
1.75	2.61	5.53	0.0232	1.73
2.00	1.51	3.21	0.0126	0.937
2.25	0.881	1.87	0.00695	0.517
2.50	0.517	1.10	0.00390	0.290
2.75	0.305	0.648	0.00222	0.165

The spurious count rate for each sensor as a function of shield thickness is then determined on the basis of the following relation:

$$C = \sum_{i=1}^{20} \xi_i \cdot \eta_i \cdot S_i \cdot A$$

where

C is the total spurious count rate, counts/s

A is the sensitive area of the detector, cm^2

S is the spectrum at the science package, photons/cm^2-s

η is the relative sensitivity, counts/photon

ξ is the relative attenuation of the shield, photons out/photon in

i (= 1 to 20) indexes the energy groups

The maximum allowable counts for each of the experiment–detector combinations is then used to interpolate for the corresponding shield thickness. Shield weights for the flat, rectangular shadow shields are then determined on the basis of these thicknesses and assumed shield areas.

5. Results

As expected, the results confirm that shielding requirements go up as one progresses from configuration 12C to 9B to 10, although there is only a factor of two spread between the 6 lb required at 18 yr by 12C and the 12 lb needed for configuration 10, assuming of course that flat, rectangular shields are used, and that scattering from support struts, etc., is negligible, allowing the modest shield areas listed in Table 1.

Table 2 tabulates spurious counts/cm^2-s versus shield thickness for each detector, assuming 18-yr-old fuel. Values for the anticipated "end of life" 12-yr-old fuel are about 5% less than for the 18-yr case, the time in which the flux due to daughter nuclides reaches a maximum. The results of the analysis are tabulated in Table 1 in which the allowable spurious count rates, minimum shield thicknesses, and weights are given. These values are for flat, rectangular shields, and can undoubtedly be reduced still further by careful shield design. Many of these values have changed a number of times as this study has progressed, and will probably continue to do so as better data becomes available. However, they do represent the best available estimates, and really substantial changes are not foreseen.

Table 3. Results for 2.17-in.-thick single shield for RTGs in configuration 12C at 18 yr

Diameter, cm	Shield weight, lb
10	16.21
11	19.61
12	23.34
13	27.39
14	31.77
15	36.47
16	41.50
17	46.85
18	54.54
19	58.52
20	64.84

Table 3 summarizes the results for a single cylindrical shield for the RTGs themselves in the case of configuration 12C. A minimum diameter of 10 cm is required to shield the fuel capsule, but larger diameters may prove necessary if scattering from the generator materials surrounding the capsule is significant. A minimum of 16 lb is required, assuming the 2.17-in. thickness needed to protect the channeltrons. The solid angle covered by the RTGs as seen from the science package in the other configurations is considerably greater and the weights that would be required to shield the RTGs in these configurations would almost certainly be prohibitive.

C. Nuclear Power Systems Definition Studies,

P. J. Gingo

1. Introduction

With the increased interest given to the possible use of plutonium dioxide radioisotope thermoelectric generators (RTG) as sources of electrical power in the NASA space science missions, it has become necessary to identify the nuclear radiations produced by an RTG and the effects produced in space science experiments which are placed in the environment of the RTG.

2. Radiation Flux

The neutron and gamma source intensities utilized in this analysis are presented in Tables 4 and 5 and are based on the production grade plutonium dioxide with the isotope composition given in Table 6. The intensity of the gamma source varies with time and is predominately due to the buildup of the daughter nuclide thallium (208). For purposes of this analysis, the radioisotope fuel was assumed to be 1000 days old.

Table 4. Neutron source intensities (nominal plutonium after aging for 1000 days)[a,b]

Neutron energy interval	Neutron source strength, neutrons/s (g PuO_2)
8.55–10.00 MeV	2.803×10^2
6.66–8.55 MeV	6.4832×10^2
5.18–6.66 MeV	1.20681×10^3
4.46–5.18 MeV	1.118363×10^3
4.04–4.46 MeV	2.224×10^3
3.14–4.04 MeV	1.19563×10^4
4.25–3.14 MeV	1.12864×10^4
1.91–2.45 MeV	6.0×10^3
1.49–1.91 MeV	2.886×10^3
1.16–1.49 MeV	1.6647×10^3
0.900–1.16 MeV	8.988×10^2
0.702–0.900 MeV	4.9887×10^2
0.546–0.702 MeV	2.8587×10^2
0.331–0.546 MeV	2.542×10^2
0.201–0.331 MeV	80.73
0.122–0.201 MeV	24.468
44.9 keV–0.122 MeV	24.63
17.0–44.9 keV	6.09
555 eV–17.0 keV	3.60
30–555 eV	0.124
3–30 eV	0.0
0.10–3 eV	0.0
0.025–0.1 eV	0.0

[a]Capsule power = 4100 W(th).

[b]Effective fuel density = 3.05 g/cm³. This is a fuel density of 10.7 g/cm³ when placed in the fuel volume of the RTG. The dilution is due to void volume and geometry considerations.

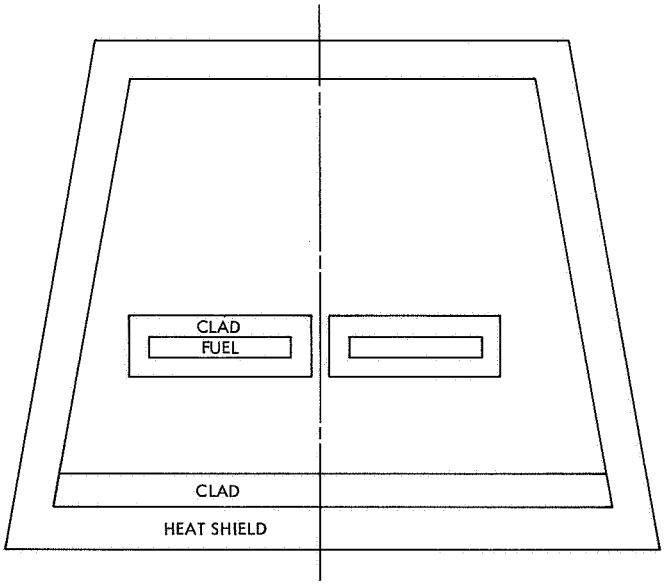


Fig. 3. Radioisotope thermoelectric generator

Table 5. Gamma source intensities (nominal plutonium after aging for 1000 days)[a,b]

Gamma energy interval	Gamma source strength, photons/s (g PuO_2)
6.0–7.0 MeV	17.84
5.0–6.0 MeV	71.07
4.0–5.0 MeV	2.126×10^2
3.0–4.0 MeV	7.015×10^2
2.0–3.0 MeV	2.464×10^5
1.8–2.0 MeV	3.203×10^4
1.6–1.8 MeV	3.233×10^4
1.4–1.6 MeV	3.416×10^3
1.2–1.4 MeV	3.111×10^4
1.0–1.2 MeV	2.925×10^4
0.9–1.0 MeV	1.162×10^3
0.8–0.9 MeV	5.704×10^5
0.7–0.8 MeV	9.608×10^5
0.6–0.7 MeV	1.702×10^3
0.5–0.6 MeV	3.264×10^4
0.4–0.5 MeV	7.625×10^3
0.3–0.4 MeV	6.619×10^4
0.2–0.3 MeV	6.161×10^5
0.044–0.2 MeV	1.723×10^8
0.001–0.044 MeV	7.198×10^8

[a]Capsule power = 4100 W(th).

[b]Effective fuel density = 3.05 g/cm³. The fuel density is defined in Table 4.

Table 6. Composition of plutonium isotopes in commercial PuO_2

Isotope	Composition, %
^{236}Pu	0.00012
^{238}Pu	81
^{239}Pu	15
^{240}Pu	2.9
^{241}Pu	0.8
^{242}Pu	0.1

The nuclear radiations were calculated for an RTG design which was selected for evaluation for the Grand Tour study, and is shown in Fig. 3. This generator contains approximately 4100 W(th) of plutonium dioxide radioactive fuel.

The calculated neutron and gamma spectra at the bottom surface (axial direction) of the RTG are given in Tables 7 and 8 respectively. The integrated (sum) neutron and gamma flux rates for detector positions exterior to the RTG are plotted in Fig. 4.

Table 7. RTG surface-radiated neutron flux[a]

Neutron energy interval	Neutron flux, number/cm²/s
10.0–8.55 MeV	2.676×10^{2}
8.55–6.66 MeV	6.334×10^{2}
6.66–5.18 MeV	1.1658×10^{3}
5.18–4.46 MeV	1.1969×10^{3}
4.46–4.04 MeV	1.8634×10^{3}
4.04–3.14 MeV	9.154×10^{3}
3.14–2.45 MeV	1.2426×10^{4}
2.45–1.91 MeV	8.24166×10^{3}
1.91–1.49 MeV	6.63255×10^{3}
1.49–1.16 MeV	4.3295×10^{3}
1.16–0.90 MeV	2.96783×10^{3}
0.90–0.702 MeV	1.95293×10^{3}
0.702–0.546 MeV	1.19117×10^{3}
0.546–0.331 MeV	1.05256×10^{3}
0.331–0.201 MeV	3.74656×10^{2}
0.201–0.122 MeV	1.4489×10^{2}
0.122–0.0449 MeV	9.2110×10^{1}
0.0449–0.017 MeV	1.88287×10^{1}
0.017 MeV–555 eV	6.7376
555–30 eV	6.033×10^{-2}
30–3 eV	1.8585×10^{-5}
3–0.1 eV	2.40×10^{-7}
0.1–0.025 eV	2.516×10^{-8}

[a]Capsule power = 4100 W(th).

Table 8. RTG surface-radiated photon flux[a]

Photon energy interval	Photon flux, number/cm²/s
7.0–6.0 MeV	3.37
6.0–5.0 MeV	1.352×10^{1}
5.0–4.0 MeV	4.065×10^{1}
4.0–3.0 MeV	1.3353×10^{2}
3.0–2.0 MeV	4.3114×10^{4}
2.0–1.8 MeV	7.225×10^{3}
1.8–1.6 MeV	7.6033×10^{3}
1.6–1.4 MeV	3.8622×10^{3}
1.4–1.2 MeV	7.960×10^{3}
1.2–1.0 MeV	8.400×10^{3}
1.0–0.9 MeV	3.187×10^{3}
0.9–0.8 MeV	5.747×10^{4}
0.8–0.7 MeV	9.376×10^{4}
0.7–0.6 MeV	3.0463×10^{4}
0.6–0.5 MeV	3.332×10^{4}
0.5–0.4 MeV	3.2820×10^{4}
0.4–0.3 MeV	3.234×10^{4}
0.3–0.2 MeV	2.679×10^{4}
0.2–0.044 MeV	2.907×10^{4}
0.044 MeV–1 keV	2.4160×10^{1}

[a]Capsule power = 4100 W (th).

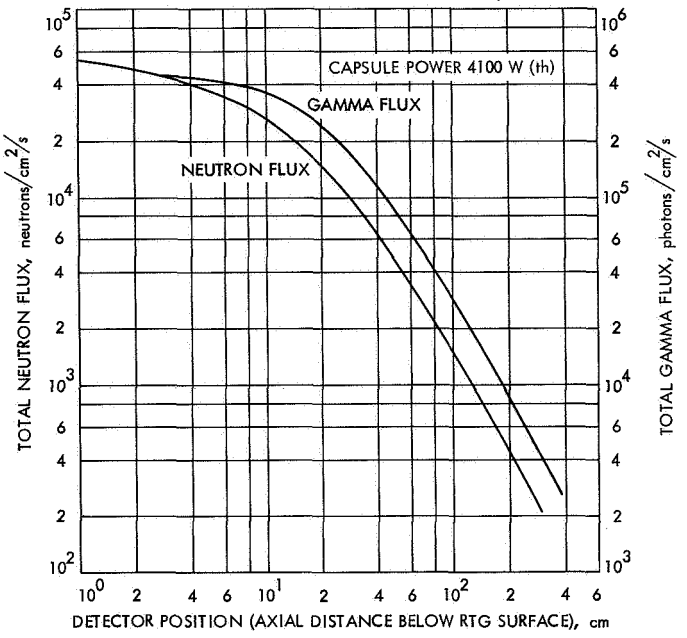


Fig. 4. Integrated flux rates vs detector position

Table 9. Attenuation ratios[a]

Photon energy interval	Tungsten thickness			
	¼ in.	½ in.	¾ in.	1 in.
7.0–6.0 MeV	0.488	0.247	0.125	0.063
6.0–5.0 MeV	0.502	0.200	0.135	0.070
5.0–4.0 MeV	0.529	0.276	0.144	0.075
4.0–3.0 MeV	0.533	0.291	0.170	0.087
3.0–2.0 MeV	0.551	0.310	0.175	0.098
2.0–1.8 MeV	0.579	0.336	0.195	0.113
1.8–1.6 MeV	0.590	0.346	0.203	0.119
1.6–1.4 MeV	0.776	0.529	0.360	0.245
1.4–1.2 MeV	0.600	0.354	0.209	0.123
1.2–1.0 MeV	0.620	0.351	0.198	0.112
1.0–0.9 MeV	0.694	0.439	0.277	0.175
0.9–0.8 MeV	0.340	0.120	0.042	0.015
0.8–0.7 MeV	0.314	0.100	0.033	0.011
0.7–0.6 MeV	0.428	0.172	0.069	0.028
0.6–0.5 MeV	0.350	0.132	0.050	0.019
0.5–0.4 MeV	0.274	0.099	0.036	0.013
0.4–0.3 MeV	0.157	0.056	0.020	0.007
0.3–0.2 MeV	0.055	0.020	0.007	0.003
0.2–0.044 MeV	← Very small numbers →			
0.044 MeV–1 keV	← Very small numbers →			

[a]The attenuation ratio is the ratio of the emerging photon flux to the incident surface photon flux. The flux units are photons/cm²/s.

3. Gamma Shielding

Tungsten metal was selected for study as a characteristic high Z shield material which would attenuate gamma rays produced by a plutonium dioxide RTG. The total gamma-ray attenuation was calculated for tungsten shields considered as slabs whose thickness varied from ¼ to 1 in. The ratios of the gamma flux emerging from the shield surface to the incident gamma flux calculated for each slab thickness and gamma energy interval are given in Table 9.

4. Neutron-Induced Gamma Radiation

The neutron flux striking the gamma will be attenuated and will produce additional secondary gamma radiation. The neutron capture and consequent gamma production were calculated only with the neutron capture gamma (n, α) reaction rates because the additional $(n, x\alpha)$ reaction rates are not available.

The gamma flux rates emerging from the various tungsten slabs were divided by the total neutron flux incident on the surface of the tungsten slab. These resulting ratios are graphically presented in Fig. 5.

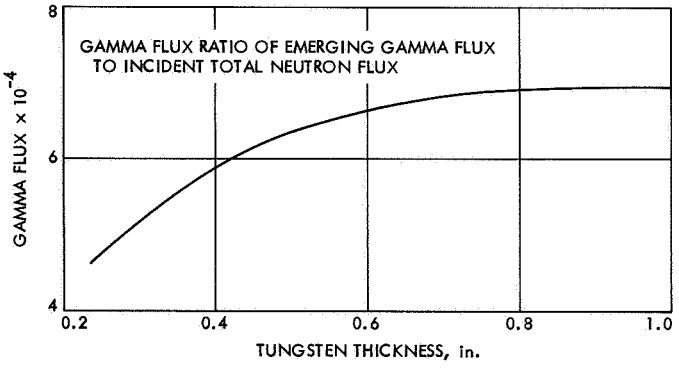


Fig. 5. Normalized gamma flux vs tungsten thickness

D. Long-Term Life Test of the Spare *Mariner* Venus 67 Power Subsystem Hardware, *A. Krug*

The advanced missions presently being considered by NASA will require power conditioning equipment (PCE) lifetimes of up to 12 yr in the missions to the outer planets. To determine the operating lifetime and performance degradation of carefully manufactured flight quality hardware, residual *Mariner* Venus 67 flight hardware will be life-tested. Failures which occur will be carefully evaluated to determine the mechanism of the failure, and to ascertain whether the failure was circuit-induced, or if the part-lifetime was exceeded. The data obtained from this test will provide a basis for the design of PCE to be used on outer planet missions.

The spare *Mariner* Venus 67 PCE will be subjected to the life test in a vacuum, with the temperature decreased over the first year (Fig. 6). The temperature deviation was chosen to simulate the worst case expected on an essentially sun-independent spacecraft bus. While the initial test period is to be approximately 1 yr, it is expected that the testing may be extended for up to 5 yr.

Performance of the PCE will be monitored, and periodic functional tests will be performed to determine the conformance of the hardware to design requirements, and to note any long-term degradation effects. At the completion of the test, the data obtained will be analyzed to identify degradation, deviations in operation, and functions or circuits which consistently exceed the requirements of the specifications.

In the event of a failure in the PCE, the testing will be terminated to allow removal of the failed part or parts, and a detailed analysis of the failure. Failed parts will be replaced and the testing continued. The failure analysis is to provide a determination of the cause; the effect of the failure on a flight mission; and recommendations for changes or improvements in the circuit design, if appropriate.

A contractor to perform this effort has been selected from seven proposers. Negotiations are presently underway, and the contractor should be started before the end of June.

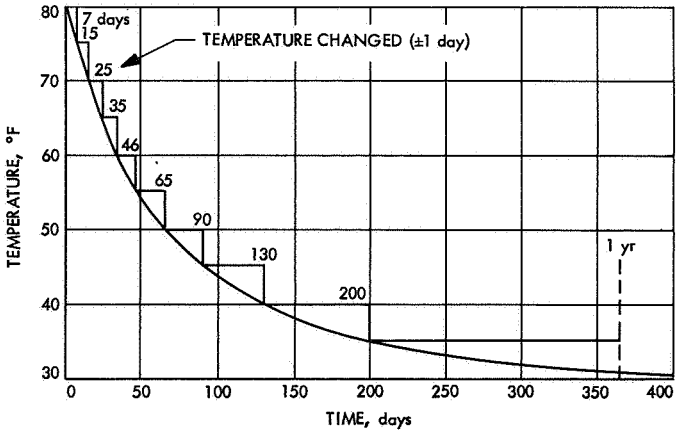


Fig. 6. Life-test temperature profile

E. Power Conditioner for a Thermionic Diode Array, *D. J. Hopper*

1. Introduction

A low-voltage converter (SPS 37-54, Vol. III, pp. 49–50) has been developed to operate from a thermionic diode array. The input voltage range is 5 to 1 Vdc, with a current range of 0 to 600 A. To handle this high current efficiently, low-saturation drop transistors developed for JPL on an advanced development contract were used. Slow rise time is the only limitation noted.

If the output of the converter becomes shorted, the current in the converter rises and will only be limited by the current capability of the thermionic diode source. However, a Hall effect device located in the positive input power line produces a voltage proportional to the input current to the converter. If the input current rises to 600 A, a level detector that senses the Hall device voltage will shut down the converter and cause a relay to close. When the relay contracts close, an external short is placed across the output of the thermionic array.

2. Test Results

The fabrication of the converter has been completed (Fig. 7). Checkout tests were performed and the unit was delivered to the Propulsion Division, which is conducting the diode experiments. The converter was exercised using the diode array as the source. The test results appear in Table 10. The efficiency for the converter is considered quite good when the source voltage and input current are taken into consideration.

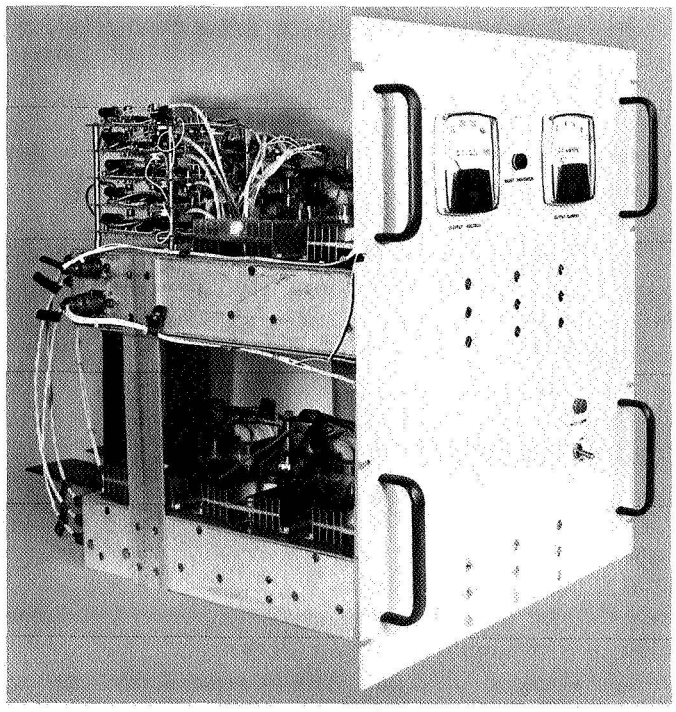


Fig. 7. Low-voltage converter

Table 10. Converter test results

V_{in}, V	I_{in}, A	V_{out}, V	I_{out}, A	Efficiency, %
0.85	245	55	2.1	57
0.95	235	68	2.1	63
1.1	220	79	1.9	69
1.25	272	95	2.4	67
1.35	295	105	2.6	69
1.4	307	110	2.8	73

A short was intentionally placed on the output of the converter, and the input current was raised slowly. At 600 A, the expected took place. The converter shut off and the short was activated. With these test results, it was concluded that the converter was operating as designed.

F. TWT Power Conditioning for Planetary Missions, *D. J. Hopper*

This task is for the design and development of a dc–dc converter for a 100-W communications traveling-wave tube (TWT) transmitter. Because of the state-of-the-art capabilities of high-power TWTs, high efficiency with high-voltage regulation is required in the dc–dc converters.

Figure 8 illustrates the electrical schematic of the TWT transmitter with the required voltage supplies. The total electrical power required for operation is 200–250 W. The helix winding shown in Fig. 8 concentrates the electron beam in the TWT and provides the RF output signal. The RF characteristics are dependent upon the steady-state value of the helix voltage, as illustrated by Fig. 9. The abscissa represents the RF power input to the TWT and the ordinate is the RF power output of the tube. Temperature variations in the exciter stage will cause variations in the RF power P_i delivered to the TWT. The helix voltage is selected to give maximum output power and not allow temperature variations in the exciter to adversely affect the RF output. As the helix voltage decreases V_{HV_2} from the nominal V_{HV_1}, the power output decreases, reducing the tube efficiency. If the helix voltage is higher than nominal V_{HV_3}, large swings in RF output power can be caused by small swings in RF input power.

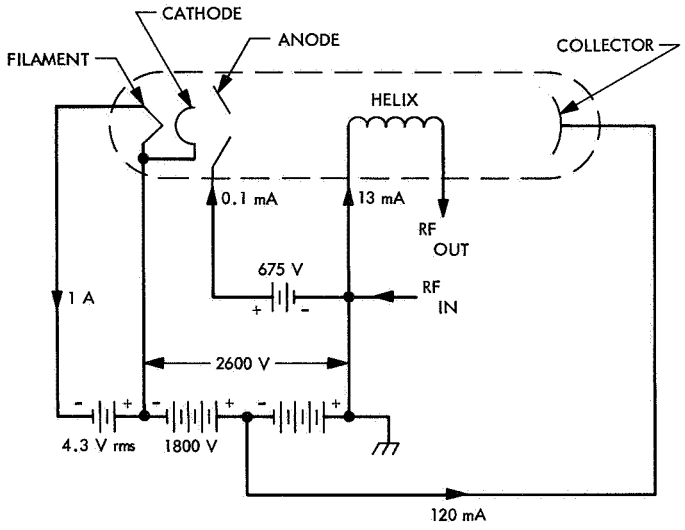


Fig. 8. TWT voltage requirement

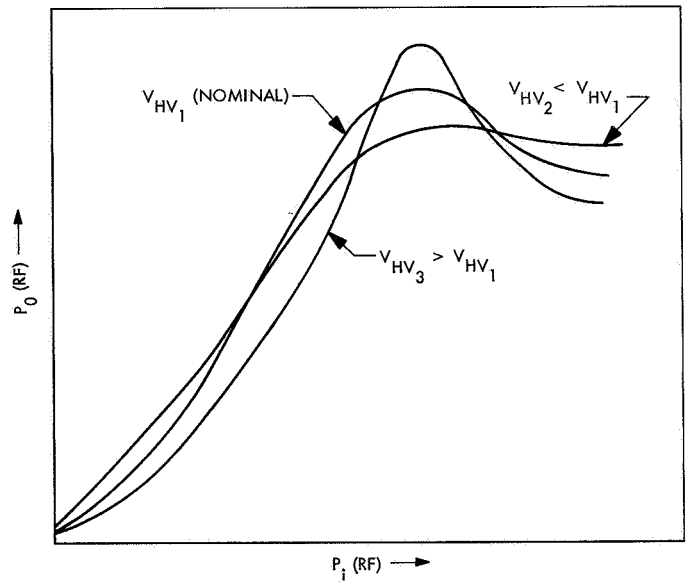


Fig. 9. TWT output characteristics

The helix voltage, then, is a critical value and must be maintained at its value with close regulation in the dc–dc converter. The design goal for regulation of the helix voltage is 0.1% with line and temperature variation. In addition, any ripple in the helix voltage will cause phase shift in the RF output. TWT specifications and the amount of acceptable phase shift determine the amount of ripple that can be tolerated on the helix. In this case, 1.0-V peak-to-peak ripple on the helix is all that can be tolerated.

For longer life of the TWT, the filament should be preheated for 90 s before the rest of the voltages are applied. Another feature desired in the converter is the ability to operate the TWT in a half-power mode. For this mode of operation, the voltages from the converter must change as follows:

Component	Half-power voltage, V	Change in voltage, V
Anode	65	210
Helix	2350	250
Collector	1530	300

Existing converters operate at efficiencies of 70 to 80%. The goal of this design is an efficiency of 90%.

At the present time, the preliminary design of the converter has been completed. The functional block diagram is illustrated in Fig. 10.

The boost regulator (SPS 37-53, Vol. III, pp. 56–57) raises the unregulated raw bus voltage (20 V ±10%) to 40 V ±0.1%. Since the TWT can be assumed to be a constant load, this is the only regulation that will be required.

The logic power portion of the circuit supplies low voltage (5 Vdc) to the 90-s delay and the drive logic, and 28 Vdc to the low-power mode selector. The 90-s delay prevents the drive logic from supplying drive power to the high-voltage inverters until the filaments have preheated for 90 s.

The filament inverter furnishes a 4.3-V rms square wave at 1.0 A to the filament when it is hot. The cold resistance of the filament is about one-fifth of the hot resistance. This means that initially the filament current would be five times as large as the final current. In the filament inverter, there will be current limiting to limit the initial current to twice the final current.

In the running inverter, there is a time when both of the switching transistors are *on*. This occurs when the storage plus the fall time of the *on* transistor is greater than the rise time of the *off* transistor. Since both transistors are *on* simultaneously, a short is placed across the inverter's source. This will cause a spike in the current and a ripple in the voltage. To prevent this, a notch t_n is placed in the inverter driving waveform (Fig. 11).

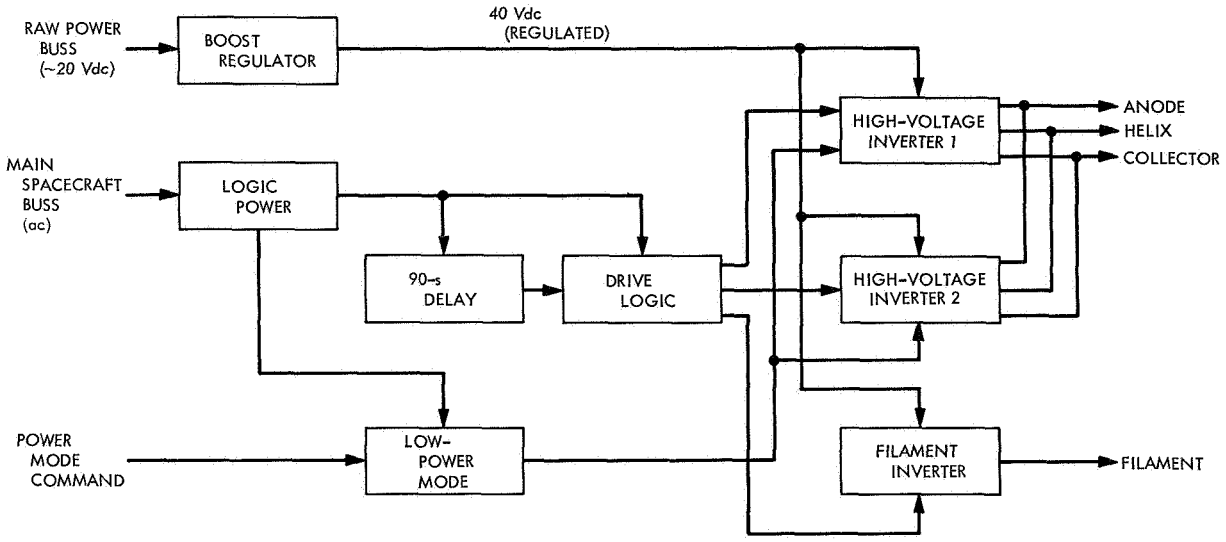


Fig. 10. TWT converter block diagram

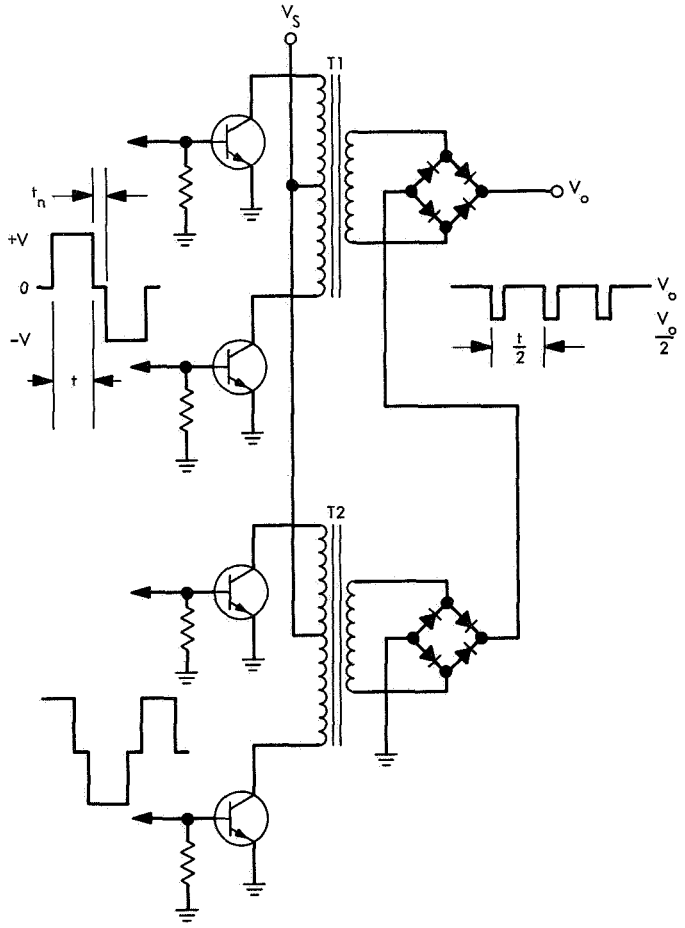


Fig. 11. TWT converter switching circuit

The notch width will be adjusted to allow the *on* transistor to turn *off* before the *off* transistor turns *on* ($t_n = t_s + t_f$). If, however, t_n is much greater than required, an undesired notch will appear in the output voltage V_o, and filtering becomes a problem. In the proposed converter, two inverters running 90 deg out of phase will be used. This means that only one inverter will be switched at a time. The output voltage, instead of falling to zero, will only fall to ½ V_o (Fig. 11). To mechanize the half-power mode, sections of the secondaries of the inverter output transformers will be switched by relays.

At the present time, the 90-s delay, the drive logic, and the filament inverter have been breadboarded, tested, and are working satisfactorily. The high-voltage section and the boost regulator have been breadboarded and are undergoing tests. Initial results show that output ripple of the high-voltage section is too high and that the boost regulator needs stabilizing. Work is in progress to correct these problems.

G. Power Conversion Circuit Development,

D. J. Hopper

The objective of the power conversion circuit development task is to develop advanced technology power conditioning components such as regulators, converters, battery chargers, etc., capable of meeting JPL advanced mission requirements.

There are several items being investigated. The first item is a high-frequency converter. This converter is being designed to operate at a frequency between 50 and 100 kHz, with an expected efficiency of approximately 90%. The input voltage is 28 Vdc, with an output voltage of 50 Vdc. The output power is 100 W. The reason for going to high frequency is to minimize weight and size. The switching losses in the transistors are being kept low by using the new high-power high-speed devices currently available. Another item being investigated is a high-frequency buck-type regulator. The running frequency of the regulator is also between 50 and 100 kHz. The efficiency design goal is 90%. An input voltage of 40 to 80 Vdc will be bucked down to 28 Vdc, with a power availability of 100 W.

A study is being made of saturation of inverter transformers (Fig. 12). If the saturation voltages of the switching transistors are different, a net dc current will flow in the transformer which may cause it to saturate. From transformer theory, it is known that the volt-seconds

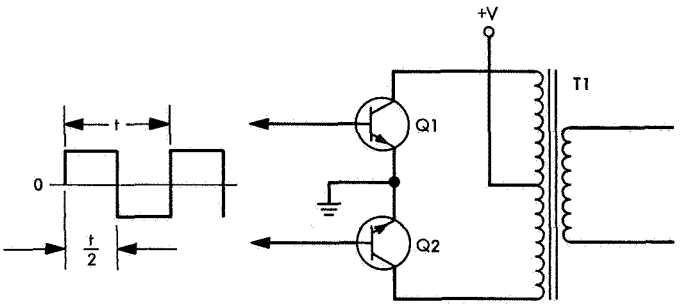


Fig. 12. Inverter transformer

when $Q1$ is *on* must equal the volt-seconds when $Q2$ is *on*. Therefore,

$$\frac{t}{2}\left[V - V_{Q1\,(sat)}\right] = \frac{t}{2}\left[V - V_{Q2\,(sat)}\right]$$

$$V - V_{Q1\,(sat)} = V - V_{Q2\,(sat)}$$

$$V_{Q1\,(sat)} = V_{Q2\,(sat)}$$

As shown in Fig. 13, more current will be forced to flow through $Q1$ to raise its saturation voltage. The dc current will be equal to

$$\frac{I_{C_{Q1}} - I_{C_{Q2}}}{2}$$

Selection of matched transistors would solve the problem, but there is no guarantee that the saturation voltage will remain the same after exposure to severe environments (radiation, temperature, etc.). The study is also attempting to determine the best transformer design to solve the problem.

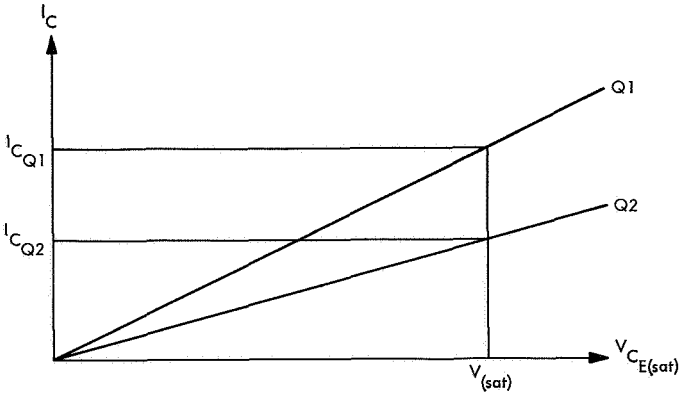


Fig. 13. Transistor characteristics

Two types of buck-boost regulators (Figs. 14a and 14b) are being investigated. The selected configuration will be breadboarded and tested. Both of these circuits operate by the same mechanism. Energy is stored in the inductor when the transistor is *on* and is transferred to the load when the transistor turns *off*. The input voltage range is 23 to 80 Vdc, and the output voltage is 56 Vdc. The output power is 400 W.

In the future, spacecraft will undoubtedly get larger. This means more power. Eventually, the source voltage will have to be increased to keep the current levels tolerable. High voltages in the order of several hundred volts would probably require the use of silicon-controlled rectifiers. Regulators using silicon-controlled rectifiers will be investigated.

Transformer core losses are higher for a square wave than for a sine wave. One of the tasks will be to compare the different waveform losses over a range of frequencies for different types of core material.

From time to time, new devices are available in industry. It is also the intent of this task to search out and investigate any new piece part that might be used in power conditioning equipment.

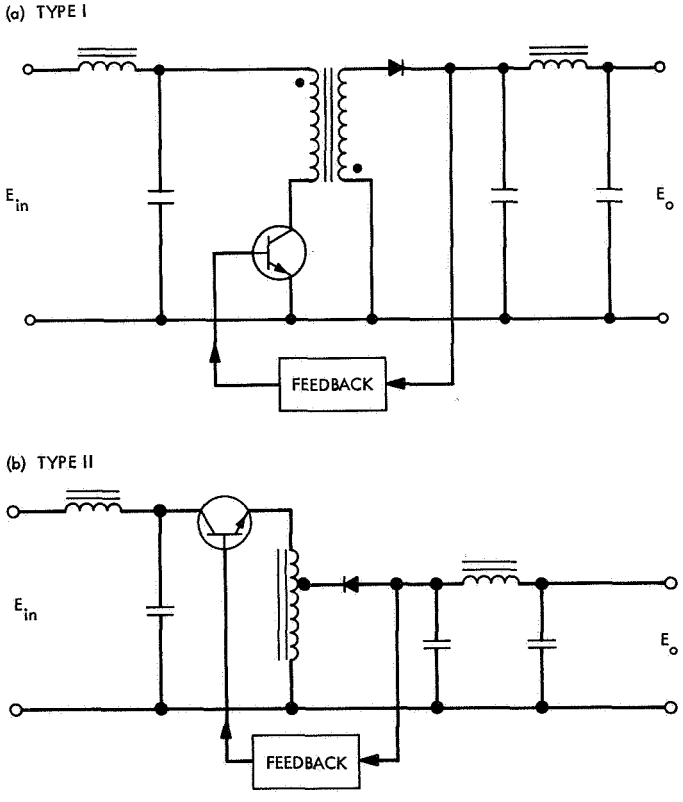


Fig. 14. Buck-boost regulators

H. Solar Power System Definition Studies, *H. Wick*

1. Introduction

During all stages of spacecraft development, it is necessary to know how much electrical power is required for each time period of the spacecraft mission. The power requirements of the spacecraft change as subsystem designs and mission objectives change. It is not practical to postpone the beginning of power subsystem design until all other spacecraft subsystems are fixed with respect to power demand and the final mission objectives are established. Thus, it becomes necessary to iterate spacecraft electrical power requirements frequently to expedite the design and sizing of the power subsystem. A power profile computer program has been developed to aid in the analysis of power requirements for the *Mariner* Mars 1971 spacecraft.

2. Computer Program Description

A simplified block diagram of the *Mariner* Mars 1971 power subsystem (Fig. 15) shows the mechanization of the power subsystem and how the various spacecraft loads are connected. Table 11 identifies the nomenclature used in the computer program. Table 12 lists the equations that are employed by the computer program. Data from the first page of the computer printout are shown in Table 13, which lists each spacecraft load as a function of the fourteen flight phases. In addition, the special and abnormal loads used by the computer program are tabulated. The special and abnormal load values are used to replace certain normal flight loads whenever the program is analyzing an abnormal gyro *on* condition, a roll control mode, or a condition where certain thermal control loads are operating from battery power.

Table 14, derived from the second page of the computer printout, lists the power versus efficiency data for each inverter, converter, and regulator of the power subsystem. This information is used by the computer to define the operating efficiency with respect to load and environmental conditions. A user option has been included in the program to allow selection of either linear or quadratic interpolation at the time of execution.

Power subsystem operational modes have been considered by the program to expedite analysis of various

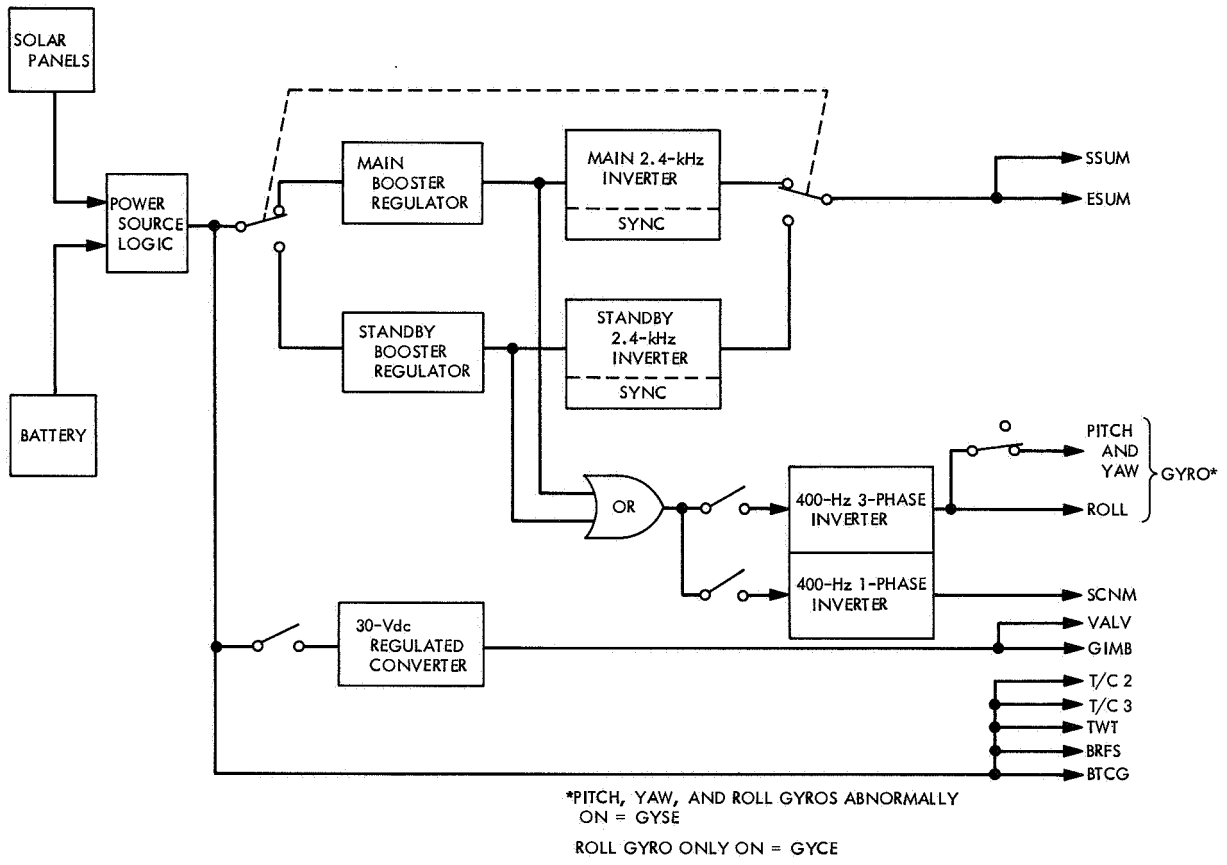


Fig. 15. *Mariner* Mars 1971 power subsystem block diagram

Table 11. Nomenclature

A/C	attitude control and autopilot	GYSO	GYRO all gyros *on*
A/CR	A/C roll control	IRIS	infrared interferometer spectrometer
A/CS	A/C all gyros *on*	IRR	infrared radiometer
BR	booster regulator	PSL	power switching and logic
BRFS	booster regulator failure sensor	PWR	power subsystem
BTCG	battery charger	PWRD	power subsystem distribution
CC&S	central computer and sequencer	PYRO	pyrotechnics
DAS	data automation subsystem	RFS	radio frequency subsystem (except TWT)
DSS	data storage subsystem	SCNE	scan control electronics
EFL	efficiency of power switching and logic	SCNM	scan control motor
EFF	efficiency	T/C1	science instruments heater
FCS	flight command subsystem	T/C2	science instruments dc heater
FTS	flight telemetry subsystem	T/C3	propulsion module dc heater
GIMB	propulsion motor gimbal	TCS2	T/C2 battery
GYCE	GYRE roll control	TCS3	T/C3 battery
GYCO	GYRO roll control	TVS	television subsystem
GYRE	gyro electronics	TWT	traveling-wave tube amplifier
GYRO	gyro motor	UVS	ultraviolet spectrometer
GYSE	GYRE all gyros *on*	VALV	propulsion motor valve

Table 12. Equations for power profile computer program

2.4-kHz inverter total science power (SSUM) = DAS + TVS + IRIS + IRR + UVS
2.4-kHz inverter total engineering power (ESUM) =
FTS + FCS + CC&S + PYRO + PWRD + DSS + A/C + SCNE + RFS + GYRE + T/C1
2.4-kHz inverter output power (TP1) = ESUM + SSUM
400-Hz (1-phase) inverter output power (TP4) = SCNM
400-Hz (3-phase) inverter output power (TP3) = GYRO
2.4-kHz inverter input power (PIN1) = TP1/EFF1
400-Hz (1-phase) inverter input power (PIN4) = TP4/EFF3
400-Hz (3-phase) inverter input power (PIN3) = TP3/EFF2
Total booster regulator output power (TBR) = PIN1 + PIN4 + PIN3
Regulated converter (30-Vdc) output power (TP11) = VALV + GIMB
Total booster regulator input power (PIN*) = TBR/EFF*
Regulated converter (30-Vdc) input power (PIN11) = TP11/EFF11
Total power switching and logic output power (PSL) = PIN* + BTCG + TWT + T/C2 + T/C3 + BRFS
Total power switching and logic input power (TOTD) = PSL/EFL
Total power demand (PSI) = PSI + PIN11

Table 13. Typical program inputs for power profile computer program

Power subsystem component	Program input for normal flight phase, W													
	Launch	Star acquisition	Cruise I (battery charge on)	Cruise II (battery charge off)	Maneuver	Preorbit TV	Orbit insertion	Playback (preorbit TV)	Orbit trim	Orbit cruise	TV sequence	Earth occultation	Canopus occultation	Playback
Science														
DAS	0.00	0.00	0.00	0.00	0.00	23.00	0.00	0.00	0.00	0.00	23.00	23.00	23.00	23.00
TVS	0.00	0.00	0.00	0.00	0.00	30.00	0.00	0.00	0.00	0.00	30.00	30.00	30.00	30.00
IRIS	2.24	2.24	2.24	2.24	2.24	21.00	2.24	2.24	2.24	2.24	21.00	21.00	21.00	21.00
IRR	0.00	0.00	0.00	0.00	0.00	2.50	0.00	0.00	0.00	0.00	2.50	2.50	2.50	2.50
UVS	0.00	0.00	0.00	0.00	0.00	14.75	0.00	0.00	0.00	0.00	14.75	14.75	14.75	14.75
Engineering														
FTS	15.20	15.20	15.20	15.20	15.20	15.20	15.20	15.20	15.20	15.20	15.20	15.20	15.20	15.20
FCS	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00
CC&S	40.50	20.00	20.00	20.00	20.50	20.00	20.50	20.00	20.50	20.00	20.00	20.00	20.00	20.00
PYRO	0.00	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25
PWRD	7.00	7.00	7.00	7.00	7.00	7.00	7.00	7.00	7.00	7.00	7.00	7.00	7.00	7.00
DSS	0.00	0.00	0.00	0.00	0.00	21.50	0.00	19.50	0.00	0.00	21.50	15.50	19.50	19.50
A/C	37.30	17.10	14.80	14.80	51.20	14.80	51.20	14.80	51.20	16.30	16.30	16.30	27.70	16.30
SCNE	4.00	4.00	4.00	4.00	4.00	18.50	4.00	4.00	4.00	4.00	18.50	4.00	4.00	4.00
SCNM	0.00	0.00	0.00	0.00	0.00	9.40	0.00	0.00	0.00	0.00	9.40	0.00	0.00	0.00
RFS	24.00	24.00	24.00	24.00	24.00	24.00	24.00	24.00	24.00	24.00	24.00	24.00	24.00	24.00
GYRE	10.00	10.00	0.00	0.00	10.00	0.00	10.00	0.00	10.00	0.00	0.00	0.00	5.00	0.00
GYRO	9.00	9.00	0.00	0.00	9.00	0.00	9.00	0.00	9.00	0.00	0.00	0.00	3.00	0.00
T/C1	4.70	4.70	4.70	14.70	14.70	10.00	14.70	14.70	4.70	4.70	0.00	4.70	4.70	4.70
T/C2	14.50	14.50	14.50	14.50	14.50	15.20	14.50	14.50	14.50	14.50	15.20	15.20	15.20	15.20
T/C3	46.50	46.50	46.50	46.50	46.50	0.00	46.50	46.50	46.50	46.50	0.00	0.00	0.00	0.00
TWT	54.00	54.00	54.00	54.00	54.00	90.00	90.00	90.00	90.00	54.00	90.00	90.00	90.00	90.00
BTCG	0.50	95.00	95.00	35.00	0.50	0.50	0.50	35.00	0.50	95.00	0.50	35.00	35.00	35.00
BRFS	1.50	1.50	1.50	1.50	1.50	1.50	1.50	1.50	1.50	1.50	1.50	1.50	1.50	1.50
VALV	0.00	0.00	0.00	0.00	25.50	0.00	25.50	0.00	25.50	0.00	0.00	0.00	0.00	0.00
GIMB	25.00	0.00	0.00	0.00	25.00	0.00	25.00	0.00	25.00	0.00	0.00	0.00	0.00	0.00

Abnormal *gyro on* load, W:

A/CS	35.10	A/CR	27.70
GYSE	10.00	GYCE	5.00
GYSO	9.00	GYCO	3.00

Special load, W:

TCS2	7.00
TCS3	21.40

Table 14. Typical power[a] vs efficiency data for power profile computer program[b, c]

PWR1	EFF1	PWR2	EFF2	PWR3	EFF3	PWR4	EFF4	PWR5	EFF5	PWR6	EFF6	PWR7	EFF7
12.5	0.560	4	0.650	4	0.650	20	0.585	20	0.600	20	0.620	20	0.870
25	0.660	5	0.680	5	0.695	35	0.670	35	0.680	35	0.700	35	0.870
37.5	0.710	6	0.710	6	0.730	50	0.758	50	0.765	50	0.773	50	0.870
50	0.745	7	0.738	7	0.760	60	0.775	60	0.782	60	0.788	60	0.870
62.5	0.775	8	0.758	8	0.780	75	0.797	75	0.803	75	0.810	75	0.870
75	0.805	9	0.775	9	0.793	85	0.809	85	0.817	85	0.825	85	0.870
87.5	0.833	10	0.790	10	0.804	95	0.819	95	0.835	95	0.838	95	0.870
100	0.853	11	0.800	11	0.814	105	0.828	105	0.840	105	0.848	105	0.870
112.5	0.867	12	0.810	12	0.820	115	0.833	115	0.848	115	0.857	115	0.870
125	0.877	13	0.820	13	0.827	125	0.837	125	0.857	125	0.865	125	0.870
137.5	0.887	14	0.828	14	0.831	135	0.840	135	0.863	135	0.870	135	0.870
150	0.903	15	0.835	15	0.836	150	0.842	150	0.870	150	0.876	150	0.870
162.5	0.903	16	0.840	16	0.840	175	0.842	175	0.876	175	0.883	175	0.870
175	0.909	17	0.846	17	0.843	200	0.842	200	0.880	200	0.887	200	0.870
187.5	0.912	18	0.850	18	0.848	225	0.842	225	0.880	225	0.887	225	0.870
200	0.915	20	0.860	20	0.850	250	0.838	250	0.880	250	0.887	250	0.870

[a]Power values are in watts.
[b]1 = 2.4-kHz main inverter.
2 = attitude control inverter.
3 = 400-Hz single-phase inverter.
4 = 27-Vdc battery booster regulator.
5 = 37.5-Vdc solar-power booster regulator (near-earth).
6 = 44-Vdc solar-power booster regulator (near-Mars).
7 = 30-Vdc regulated converter.
[c]EFL = 0.970.

combinations of flight phases, power subsystem efficiencies, abnormal gyro *on* conditions, and roll control conditions. These modes are selected by the user at the time of execution through the use of fourteen control cards. Any or all operational modes can be computed for each flight phase.

Table 15, derived from the computer printout, shows the program output for a typical flight phase, in this example, orbit insertion. Each power supply of the power subsystem is listed with its respective loads. The power output, operating efficiency, and power input for each supply are computed and printed. This process continues automatically until the total power demand for each selected operational mode of a particular flight phase is computed. The program then advances the printout to a new page, and the process is repeated for the subsequent flight phase and modes.

Data from the last page of the computer printout are given in Table 16, which summarizes the total power demand as a function of mission flight phase and power subsystem operational mode. The *Mariner* Mars 1971 power profile computer program is written in Fortran IV and presently is run on the IBM 7094. This program has been converted for operation with the UNIVAC 1108.

Table 15. Typical output from power profile computer program for a given mission flight phase (orbit insertion)

2.4-kHz main inverter	
Average power load, W	
Science	
DAS	0.00
TVS	0.00
IRIS	2.24
IRR	0.00
UVS	0.00
Total	2.24
Engineering	
FTS	15.20
FCS	3.00
CC&S	20.50
PYRO	0.25
PWRD	7.00
DSS	0.00
A/C	51.20
SCNE	4.00
RFS	24.00
GYRE	10.00
T/CI	14.70
Total	149.85
Power output, W	152.09
Efficiency	0.903
Power input, W	168.43

Attitude control inverter	
Average power load, W	
GYRO	9.00
Power output, W	9.00
Efficiency	0.775
Power input, W	11.61

400-Hz single-phase inverter	
Average power load, W	
SCNM	0.00
Power output, W	0.00
Efficiency	0.00
Power input, W	0.00

30-Vdc regulated converter	
Average power load, W	
VALV	25.50
GIMB	25.00
Power output, W	50.50
Efficiency	0.870
Power input, W	58.05

Normal operational mode (battery)	
BR output, W	180.04
BR efficiency	0.842
BR input, W	213.82
PSL average load, W	
BTCG	0.50
T/C2	7.00
T/C3	21.40
BRFS	1.50
TWT	90.00
BR input	213.82
PSL output, W	334.22
PSL efficiency	0.970
PSL input, W	344.56
Total power demand, W	402.61

Normal operational mode (near-Mars)	
BR output, W	180.04
BR efficiency	0.884
BR input, W	203.71
PSL average load, W	
BTCG	0.50
T/C2	14.50
T/C3	46.50
BRFS	1.50
TWT	90.00
BR input	203.71
PSL output, W	356.71
PSL efficiency	0.970
PSL input, W	367.74
Total power demand, W	425.79

Table 16. Typical output of summary page from power profile computer program

Power subsystem mode	Total power demand/normal flight phase, W													
	Launch	Star acquisition	Cruise I (battery charger on)	Cruise II (battery charger off)	Maneuver	Preorbit TV	Orbit insertion	Playback (preorbit TV)	Orbit trim	Orbit cruise	TV sequence	Earth occultation	Canopus occultation	Playback
Normal (battery)	331.14	351.80	0.00	0.00	365.50	0.00	402.61	0.00	391.24	325.67	0.00	0.00	0.00	0.00
Normal (near-earth)	0.00	380.86	355.25	303.91	390.36	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Normal (near-Mars)	0.00	0.00	353.85	302.50	388.68	408.75	425.79	361.27	415.41	355.39	398.64	401.57	431.60	406.38
Gyros on (near-earth)	0.00	0.00	401.05	350.50	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gyros on (near-Mars)	0.00	0.00	0.00	349.04	0.00	457.95	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Roll control (battery)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	350.37	0.00	0.00	0.00	0.00
Roll control (near-Mars)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	385.91	0.00	378.25	423.75	426.84	0.00	431.60

I. Liquid Amalgam Electrodes, *G. L. Juvinall*

1. Introduction

The results of high-rate discharge and stand-life investigations of liquid amalgam electrodes were reported in SPS 37-55, Vol. III, pp. 105–107. The remarkable properties of these electrodes were discovered during an investigation of the reactions of electrode materials used in alkaline batteries. Dr. G. Myron Arcand of Idaho State University is the principal investigator on this project.

The results of open-circuit potential studies of sodium-amalgam electrodes are reported here. These data represent part of a major objective of the program, i.e., to explore the field of liquid amalgam electrodes from the point of view of power-producing reactions. An additional objective is to clearly define the electrochemical processes involved in their operation.

2. Experimental approach

The general experimental methods have been previously reported under this contract (Refs. 1 and 2). A number of measurements were made to determine the effects of metal concentration in the amalgam and sodium hydroxide concentration in the electrolyte solution on the open-circuit voltage. The concentration of metal in the amalgam was calculated from the coulombs of electricity passed during charge of the electrodes. The assumption was made that current efficiency is 100%. All potential measurements were made with a Dohrmann Model RSC 1100 recorder against a Hg/HgO–10 *N* KOH reference electrode.

Table 17. Effect of amalgam concentration and electrolyte concentration on the open-circuit potential of the amalgam cathode

Concentration of Na in Hg, moles/kg	E_{oc}, V vs Hg/HgO		
	5 *N* NaOH	10 *N* NaOH	Saturated NaOH
0.038		−1.72	−1.65
0.045	−1.82		
0.075		−1.75	−1.66
0.076	−1.89		
0.170	−1.89		
0.187		−1.80	−1.73
0.298	−1.91	−1.79	−1.72

3. Results and discussion

Table 17 shows the effects of both amalgam concentration and electrolyte concentration on the value of E_{oc} (open-circuit voltage). It should be observed that E_{oc} increases with increasing metal concentration in the amalgam for any fixed electrolyte concentration, and that it decreases with increasing electrolyte concentration at a fixed amalgam concentration.

If the electrode reaction is assumed to be reversible and corresponds to the equation

$$Na^+ + Hg + e' \rightarrow Na\,(Hg)$$

then the Nernst equation may be applied:

$$E = E^0 - \frac{RT}{NF}(\ln a_{amal} - \ln a_{Na})$$

Although there are no data regarding the metal activities in the amalgams, it may be seen that the directions of the observed trends are what would normally be expected. However, the magnitudes of the effects are considerably greater than expected. As seen in Table 18, doubling the concentration of NaOH causes an increase of 100 mV as compared to an expected 20-mV change. A five-fold increase in the activity coefficient of NaOH would be required in order to account for these observations. Additional work will be done in this area.

Table 18. Effect of electrolyte concentration on the open-circuit potential of the amalgam cathode

Concentration of Na in Hg, moles/kg	E_{oc}, V vs Hg/HgO		
	5 *N* NaOH	10 *N* NaOH	Saturated NaOH
0.228			−1.70 ±0.01
0.278		−1.78 ±0.02	
0.280	−1.88 ±0.00		

Earlier work has shown that stand life is quite poor when 5 *N* NaOH electrolyte is used; gassing becomes an immediate problem. However, it is in this region of electrolyte concentration that the most reproducible and negative cathodic potentials are obtained. In order to promote the development of this electrode system into a useful device, the search for a practical nonaqueous diluent which will permit better stand-life behavior will be continued.

References

1. Arcand, G. M., *Investigation of Electrode Materials for Alkaline Batteries,* First Quarterly Report, Contract 952265. Idaho State University, Pocatello, Idaho, Aug. 25, 1968.

2. Arcand, G. M., *Investigation of Electrode Materials for Alkaline Batteries,* Second Quarterly Report, Contract 952265. Idaho State University, Pocatello, Idaho, Nov. 25, 1968.

J. State-of-Charge Indicator Investigation,

S. J. Krause

1. Introduction

The need for obtaining additional data from spacecraft batteries during future missions is of great significance in view of the importance of a properly functioning battery's relation to proper spacecraft functioning and the successful completion of the mission. The most critical period of battery function in the past has been during the launch of the spacecraft, until the solar panels acquire the sun. The battery may be used for subsequent midcourse corrections, and throughout the mission, the battery is a standby source of power at all times. Future JPL missions will require an even greater degree of battery usage and participation in supporting critical spacecraft functions.

It is, therefore, essential that more data on battery condition and capacity be made available throughout the mission, including the period prior to launch. This additional data will enable engineers to predict battery performance more accurately during all phases of the mission and to evaluate general battery condition at any time.

The need for such data may in part be met by the development of a state-of-charge indicator. The value of a state-of-charge indicator cannot be over-emphasized, since it is possible that the planning and performance of critical spacecraft functions on future missions may depend upon the knowledge of available power from the battery. The prime objective of this task has been to develop techniques whereby the available energy output of a battery may be determined at any time, with a future capability of being integrated into structurally incorporated devices for use within the battery.

At the present time, emphasis has been placed on development of state-of-charge indicators and measuring techniques for silver–zinc batteries, although a great deal of effort has been expended by other investigators in developing techniques for nickel–cadmium batteries and other alkaline systems (Refs. 1–5). The following is a list of techniques which have been tried in the attempt to develop a practical and accurate state-of-charge indicator:

(1) Cell voltage phase shift (Refs. 1 and 4).

(2) Discharge and charge polarization voltage (Refs. 1 and 4).

(3) Integrating ampere-hour meter (Refs. 1, 4, 5, and 6).

(4) Discharge and charge voltage (Ref. 6).

(5) Double-layer capacitance (Ref. 2).

(6) Transient peak height (Ref. 2).

(7) Audio-frequency impedance (Ref. 2).

(8) Internal resistance (Ref. 2).

(9) Colorimetry (Ref. 2).

(10) Resonant frequency (Ref. 2).

(11) Rate of oxygen recombination (Ref. 3).

(12) Charge acceptance rate (Ref. 6).

(13) Battery temperature (Ref. 6).

The general conclusion which may be drawn from all work based on the listed techniques and others which have not been formally published is that all are unsuccessful in accurately determining the state-of-charge of a battery. The effort at JPL has been directed toward a technique which appears to have been as yet untried, but shows some promise in its accuracy of charge-state determination and feasibility for being structurally integrated into the battery.

2. Approach and Results

Each compound used in the active material of a battery electrode has infrared absorption bands peculiar to that compound. Different compounds have different absorption bands at different wavelengths in the infrared region of the electromagnetic spectrum. The presence of a given compound decreases the intensity of transmitted radiation at its related absorptive wavelength by

$$I = I_0 e^{-\alpha x}$$

where

I = transmitted intensity

I_0 = source intensity

α = absorption coefficient

χ = amount of absorbing material

It was postulated that measuring the transmission of infrared radiation through a part or parts of the active material in a silver battery electrode could provide a quantitative determination of the charge and discharge associated compounds, including AgO, Ag_2O, and elemental silver (Ag). As a preliminary step, a review of the literature was completed to better define the infrared absorption characteristics of the active materials. Ag_2O was found to have peaks at 7, 7.4, 9.3, and 12.5 μm (Ref. 7). AgO, in the pure state, has some peaks at 10 and 14 μm (Ref. 8). Information, for possible future work on the negative active material, was also obtained showing that ZnO had a wideband absorption peak at 5.5 (Ref. 9), 6.3 (Ref. 10), 10.1 (Ref. 11), and 11.5 μm (Ref. 11).

As a result of the findings in the literature, the investigation of the absorption characteristics of the positive electrode active materials (AgO, Ag_2O, Ag) was limited to the 2.5- to 15-μm infrared region. Spectrophotometric scans of small parts of electrode structures were performed on a Perkin–Elmer 421 recording spectrophotometer. The results of the measurements indicated that very high light intensities would be needed to achieve a usable degree of transmission through the electrode material when a great deal of the relatively opaque Ag was present. However, this approach was not determined to be impracticable, since it had the capability of being incorporated into a battery. The integration of a photoemitter and photodetector, with a small separation gap of approximately 100 μm, into the positive electrodes is contemplated. Small devices of relatively high intensities at the selected infrared wavelength are available if such a technique could correlate transmitted intensity with the quantitative amount of an absorbing compound such as AgO or Ag_2O.

An alternative approach, that could surmount the difficulty of obtaining sufficient transmission when the radiant energy is passed through the material, is to use a "light pipe" incorporated into a battery, and to use attenuated total reflectance (ATR) spectroscopy as a means of determining the feasibility of the method.[4]

If light is incident on a plane interface of two optical media at an angle exceeding the critical angle, the interface is 100% reflecting. It is common in "light pipe arrangements" to have 10^2 reflections, after which 50% of the input power is measured.[4] This gives a reflectance of $X^{100} = 0.5$ where

$$\log X^{100} = \log 0.5$$

$$\log X = -0.00301$$

$$X = 0.994$$

and, therefore, high reflectances are obtained. This light does have the ability to penetrate the interface to a small depth, as shown in Fig. 16. This behavior is also predicted by Maxwell's theory[4] by the following equation:

$$d_{1/2} = \frac{0.693\,\lambda}{2\pi\left(\sin^2\theta - n_{21}^2\right)^{1/2}}$$

where

$d_{1/2}$ = depth of penetration at 1/2 power

θ = angle of incidence

$$n_{21} = \frac{\text{index of refraction of rare medium}}{\text{index of refraction of dense medium}}$$

λ = wavelength

It is not difficult to obtain penetration up to λ.

[4]Jensen, R. J., Brigham Young University, Provo, Utah (private communication).

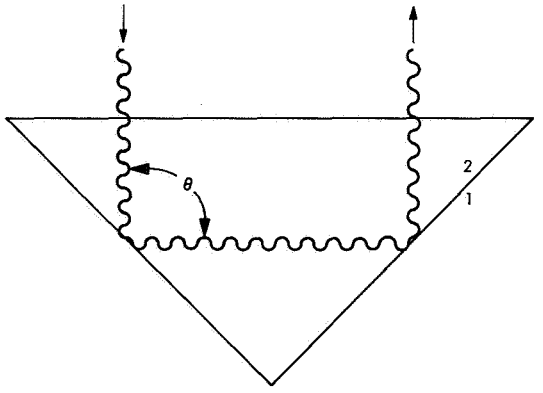


Fig. 16. Radiation penetration at planes of attenuated total reflections (1 rarer medium, 2 denser medium)

If the material of the rarer medium absorbs at the wavelength in question, absorption of the "totally" reflected beam takes place to an extent comparable to the absorption that would take place if the beam had traversed a thickness of the rarer material, in this case, the electrode active material. The beam is attenuated at those wavelengths where the material absorbs, with the attenuation level occurring as a function of the quantity of absorbing material present at the interface. Given sufficient surface area to the light pipe, as described in a possible geometry shown in Fig. 17, a representative sampling of the electrode active material could be measured and, thus, the state-of-charge might be determined.

A series of tests was performed on standard silver electrodes at various states of charge and discharge. Spectrophotometric scans were made on a Perkin–Elmer 421 recording spectrophotometer, combined with a Wilkes model 9 internal reflecting attenuated total reflectance attachment.

The results, a sampling of which may be seen in Fig. 18, were of great interest, since a definite difference could be seen in attenuation of the light beam by the electrodes in different states. It should be noted, however, that the absorption was greater for an uncharged

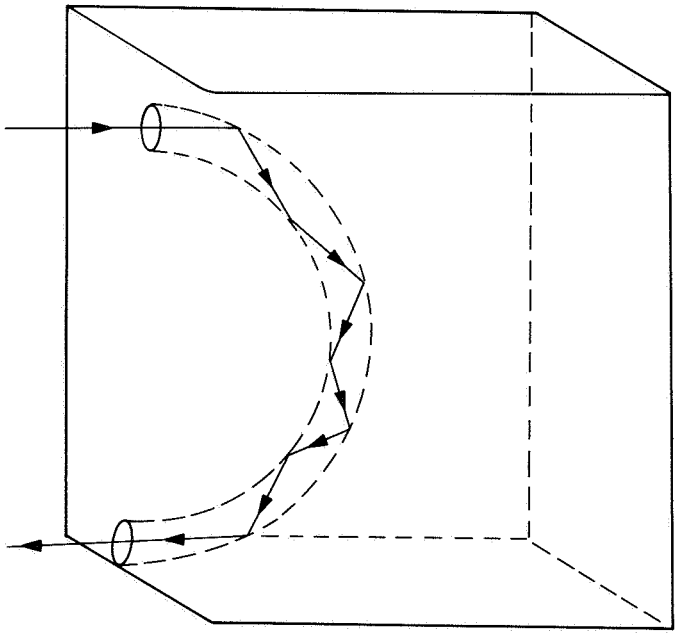


Fig. 17. Path of infrared radiation traveling through silver electrode via light pipe

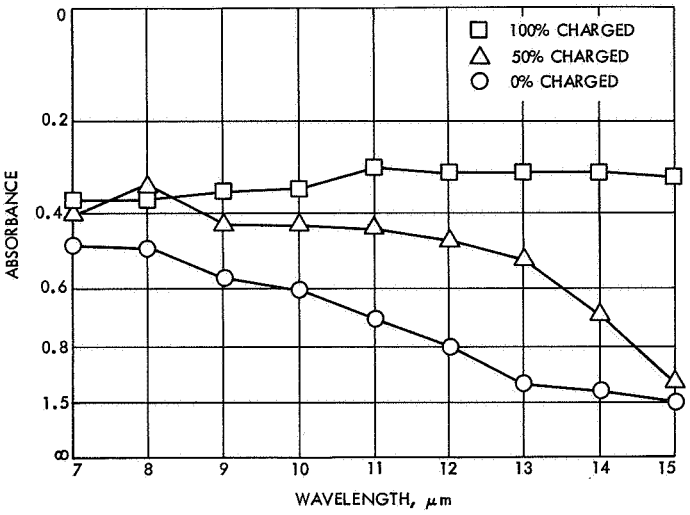


Fig. 18. Standard silver electrode wavelength vs absorbance

electrode, and seemed to decrease as the charge state increased. This is not the expected result, since more absorption should occur when the charge state is higher, there being more AgO and Ag_2O present.

3. Conclusions

Based on the results of the ATR testing, a trend has been established between infrared absorption of light in the 12- to 15-μm region and the charge state of standard silver electrodes. However, since this trend was the reverse of that which would normally be expected, it could be suggested that the attenuation of the light was a function of particle size differences, causing differences in scatter. The active material particle size and grain structure change as an electrode is charged and discharged (Ref. 12). This, of course, does not preclude further pursuit of this technique, since an unusual but useful trend has been established, but also because other applications of such a technique may exist. For example, it may be possible to use such an effect to nondestructively test silver electrodes for particle size and density as they are sintered, thereby increasing the uniformity of silver electrodes fabricated into batteries. The usefulness of the technique for application as a state-of-charge indicator must, of course, be explored further before any definite conclusions can be drawn. No work has as yet been performed on the negative electrode active material, which may also show promise. A trend has been observed which does indicate that there could be many uses and applications for infrared measurements in studies of electrode materials, limited only by the amount of future effort expended in investigations.

References

1. Cherdak, A., and Shair, R., "Determination of State-of-Charge on Ni-Cd batteries," in *15th Annual Proceedings of the Power Sources Conference,* May 9, 1961.

2. John, F. J., *Non-Destructive State-of-Charge Tests for Batteries.* U.S. Army Electronics Research Laboratory, Ft. Monmouth, N.J.

3. Turner, D. R., *The Effect of State-of-Charge of the Cadmium Electrode on Oxygen Recombination in Sealed Ni-Cd Cells.* Bell Telephone Laboratories, Inc., Murray Hill, N.J.

4. Lurie, H., Seiger, and Shair, R., *State-of-Charge Methods for Ni-Cd Batteries.* Gulton Industries, Inc., Metuchen, N.J.

5. *State-of-Charge Indicators for Alkaline Batteries,* 3rd Quarterly Technical Progress Report. Gulton Industries, Inc., Metuchen, N.J., Nov. 15, 1962.

6. Tucker, W. M., *Spacecraft Battery Charge Control and Cell Monitoring.* NASA Goddard Spaceflight Center, Greenbelt, Md.

7. Kreingold, F. I., "Optical Absorption of Ag_2O in the Infrared," *Opt. Spektrosk.,* Vol. 20, No. 2, pp. 336–337, 1966.

8. Banjerjee, R. S., and Baser, S., "Ligand Field Spectra of Some Metals in Their Unfamiliar Oxidation States," *J. Inorg. Nucl. Chem.,* Vol. 26, No. 5, pp. 821–824, 1964.

9. Arneth, R., "Absorption of ZnO Single Crystals in the Infrared," *Z. Phys.,* Vol. 155, pp. 595–608, 1959.

10. Matsushita, S., and Nakata, T., "Infrared Absorption of Zinc Oxide and of Absorbed CO_2," *J. Chem. Phys.,* Vol. 36, No. 3, pp. 665–667, 1962.

11. Miloslavskii, V., and Kovalenko, N., "Absorption of ZnO in the Infrared Region of the Spectrum," *Opt. Spektrosk.,* Vol. 5, pp. 614–617, 1958.

12. Wales, C. P., and Simon, A. C., *Feasibility of Microscopy for Investigating the Silver Oxide Electrode,* Report No. 6647. Naval Research Laboratory, Washington, D.C., Jan. 9, 1968.

X. Spacecraft Control

GUIDANCE AND CONTROL DIVISION

A. Spurious Canopus Tracker Outputs, *R. H. Burns*

1. Introduction

The *Mariner* Mars 1969 Canopus tracker is being studied to evaluate the design for use on outer-planet missions, where long-term reliability and fully automatic operation will be required. One recently completed task was the study of spontaneous roll error outputs that have been seen with several of the flight model trackers during laboratory testing. The anomalous operation has occurred when the tracker is in a backup mode with the star recognition logic disabled via ground command, so that the tracking circuits are free to track stars of any intensity. When a tracker in this uninhibited mode is not looking at a star, as would happen during a roll search, its field of view is normally biased hard to one side. Sometimes, however, the field of view has been seen to leave this position and wander about at random.

The effects of such an event during a roll search could be serious. Wandering of the roll error signal would result in a random firing of the control jets, with no assurance that the spacecraft would perform a systematic roll search.

2. Normal Operation of the Tracking Circuits

Figure 1 is a functional diagram of the *Mariner* Mars 1969 Canopus tracker signal processing circuitry with the recognition logic disabled. The dissector tube is an electrostatically scannable photomultiplier. Its field of view can be biased anywhere within a 9° arc that is roughly in the plane of spacecraft roll (perpendicular to the sun line). There is a small correction from the roll plane to allow for Canopus' angular displacement from the south ecliptic pole. Within the field of view, the tube makes a 1200-Hz sawtooth scan, so that the output is approximately a 1200-Hz sine wave whose phase indicates the position of the observed star within the field of view. The demodulator measures this phase and generates an error signal that is integrated to produce the dc roll bias. Thus, the dc bias is constantly corrected so that the field of view is centered upon the star. The tracker roll angle error signal is proportional to the dc roll bias.

The intensity loop is a selective gain control that varies the dissector tube dynode voltage in such a way that the 1200-Hz tube output component is held constant for a wide range of optical input intensities. The tracker intensity output is derived from the dynode voltage.

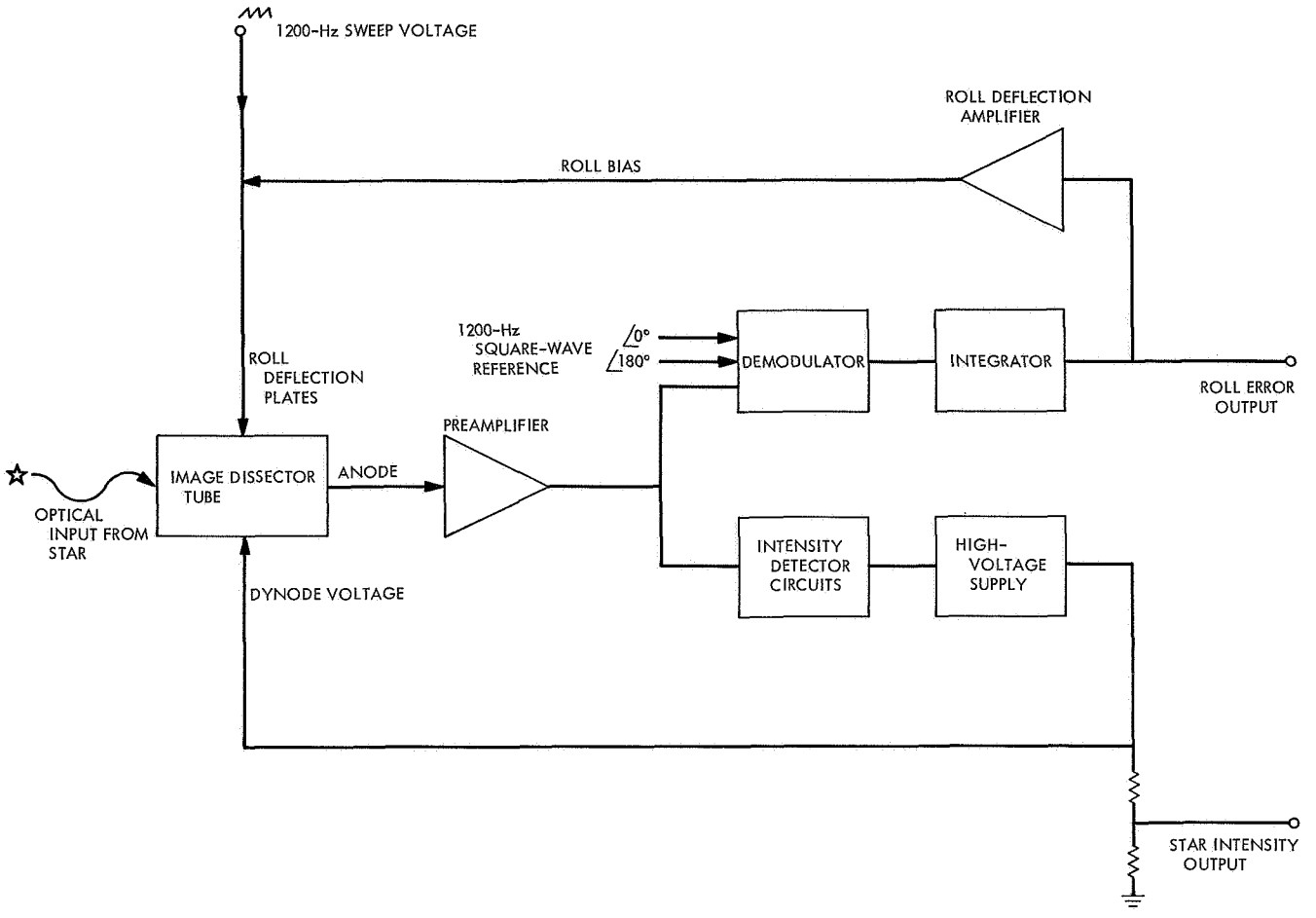


Fig. 1. Canopus tracker signal processing

3. Spurious Output Mode

The signal from the S-11 dissector tube is quite noisy. When the dynode multiplier gain is varied by change in the dynode voltage, the rms noise magnitude remains constant with respect to the gain. The noise, therefore, is suspected to come from the cathode or the first dynode of the multiplier structure. In some cases, all of this noise may be attributed to the predicted cathode shot noise. Many tubes, however, are much noisier than the shot noise equation predicts and lead to spurious tracker roll error outputs. Since disconnecting the cathodes of these tubes has very little effect on the noise level, the excess noise source is thought to be at the first dynode rather than at the cathode. The source is suspected to be some sort of breakdown mechanism involving cesium vapor from the S-11 photocathode or the dynodes.

A little understood process known as "spot-knocking," which consists of applying 4000 V to the dynodes for several hours, seems to reduce the extra noise. If the noise is a cesium vapor phenomenon, "spot-knocking" may clean out the tube by ion pumping. Experience with a recently designed photomultiplier tube shows that the excess noise can in some cases be eliminated.

The demodulator is shown in Fig. 2. The switches are field-effect transistors driven by square waves in phase

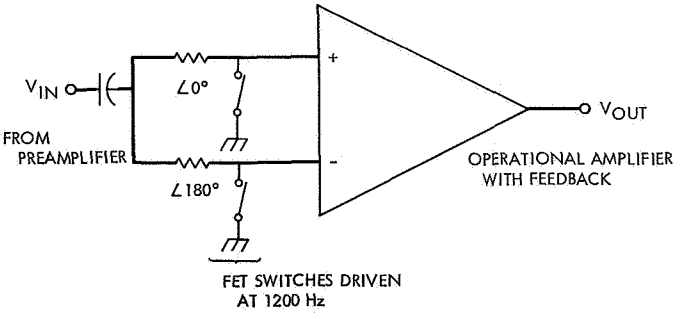


Fig. 2. Demodulator circuit

with the dissector tube sweep voltage. The circuit output consists of low-pass noise plus a dc component that is

$$V_{out} = 1.3\,V_{in} \sin\theta$$

where θ is the electrical phase angle between the star signal from the dissector tube and midpoint of the sawtooth wave sweep. V_{in}, the zero-to-peak magnitude of the star signal, is normally regulated to 5 V or less by the intensity loop. If the tube is extremely noisy, however, the peak value of the signal plus noise fed to the demodulator can reach the output saturation level of the preamplifier, which is limited to 15 V by the preamplifier power supply. The demodulator also uses a 15-V supply, but in the *Mariner* Mars 1969 tracker configuration, the demodulator has a gain of 1.3, so that the corresponding modulator input voltage is about 11 V. Thus, the preamplifier can saturate the demodulator. Since the demodulator can not have exactly equal positive and negative saturation voltages, a signal large enough to saturate it can produce a dc output component. If the recognition logic is disabled, this output can cause the integrator to drift at random, as has been observed.

Prevention of this phenomenon requires that the preamplifier be incapable of saturating the demodulator. The maximum allowable demodulator input may be increased to the maximum preamplifier output by reducing the demodulator gain to unity. In order to preserve the same loop gain, the gain of the following integrator stage should be raised by the same factor by which the demodulator gain is reduced.

Further protection from demodulator saturation may be had by limiting the maximum available preamplifier output. In the current model tracker, the preamplifier is limited by 10-V zener diodes, but is designed in such a way that there is still some gain above this output voltage. Large noise spikes, although attenuated somewhat, may then pass on to the demodulator. Noise resistance can be increased by using hard limiting at a voltage below the maximum allowable demodulator input level.

A test was made with a breadboard tracker to determine the effects of hard preamplifier limiting and reduced demodulator gain. Before the modifications were made, sufficient random noise was injected into the preamplifier to fully saturate it. The excursions of the integrator were only limited by its saturation voltages, i.e., 30 V, peak to peak. After the demodulator gain was lowered to slightly less than unity and the preamplifier hard limited at 12 V, the same level input integrated to about 30 mV peak. This corresponds to an error of 8×10^{-3} deg of star null offset.

4. Conclusion

Although trouble has been experienced during testing with spurious roll error outputs from the *Mariner* Mars 1969 Canopus tracker, the problems do not seem unavoidable. The difficulties seem to be caused by saturation of the demodulator circuit by noise from the dissector tube. The sensitivity of the circuitry to this noise may be drastically reduced by proper selection of demodulator gain and preamplifier saturation level.

B. Nozzle Thrust at Leakage Flow Rates,

F. G. Roselli-Lorenzini

The attitude-control thrusting of the *Mariner* spacecraft is performed by a nitrogen cold gas system whose torque-producing components are 12 solenoid-actuated jet valves, four for each axis and two for each axis direction.

The specification leak rate of a single *Mariner* valve is 3 cm^3/h measured at standard conditions with a 15-psig pressure drop across the valve seat. In vacuum, with the same pressure drop, the measured leak rate will decrease by a factor of about 2, due to the lesser density of the gas at the valve inlet. The gas, leaking through the valve nozzle, produces a thrust which can be detected by the spacecraft dynamics if an unbalance is present between the leak rates of the four valves of the same axis. The effects may be an increased gas expenditure to counteract the leakage torque and a spacecraft trajectory deviation.

Some attempt has been made in the past to analytically evaluate the nozzle thrust at a leakage flow rate and correlate the results with the telemetered data on the spacecraft angular position. The latest approach[1] is based on the "billiard ball" theory applied to the gas molecular flow rate through the nozzle. The result is a linear relationship between thrust and leak rate at a ratio of 0.0373 $\mu lb/cm^3/h$. As a proof check for the analytical results, a test program has been established to measure directly, in vacuum, the nozzle thrust using spare valves from the *Mariner IV* and *V* cold gas systems. A valve cutaway view is shown in Fig. 3. The thrust stand was a torsional balance in which a 0.005-in.-diam stainless-steel music wire supported the thrusting system (a 5-in.-diam pressure vessel, feeding valves, tubing, a manifold, jet valve, and nozzle) on an aluminum platform. The thrust stand was contained in a small vacuum chamber. The readout mechanism consisted of a lamp and a photocell shadowed by an arm extending from the thrust platform.

[1]By E. A. Laumann, JPL Environmental Sciences Division.

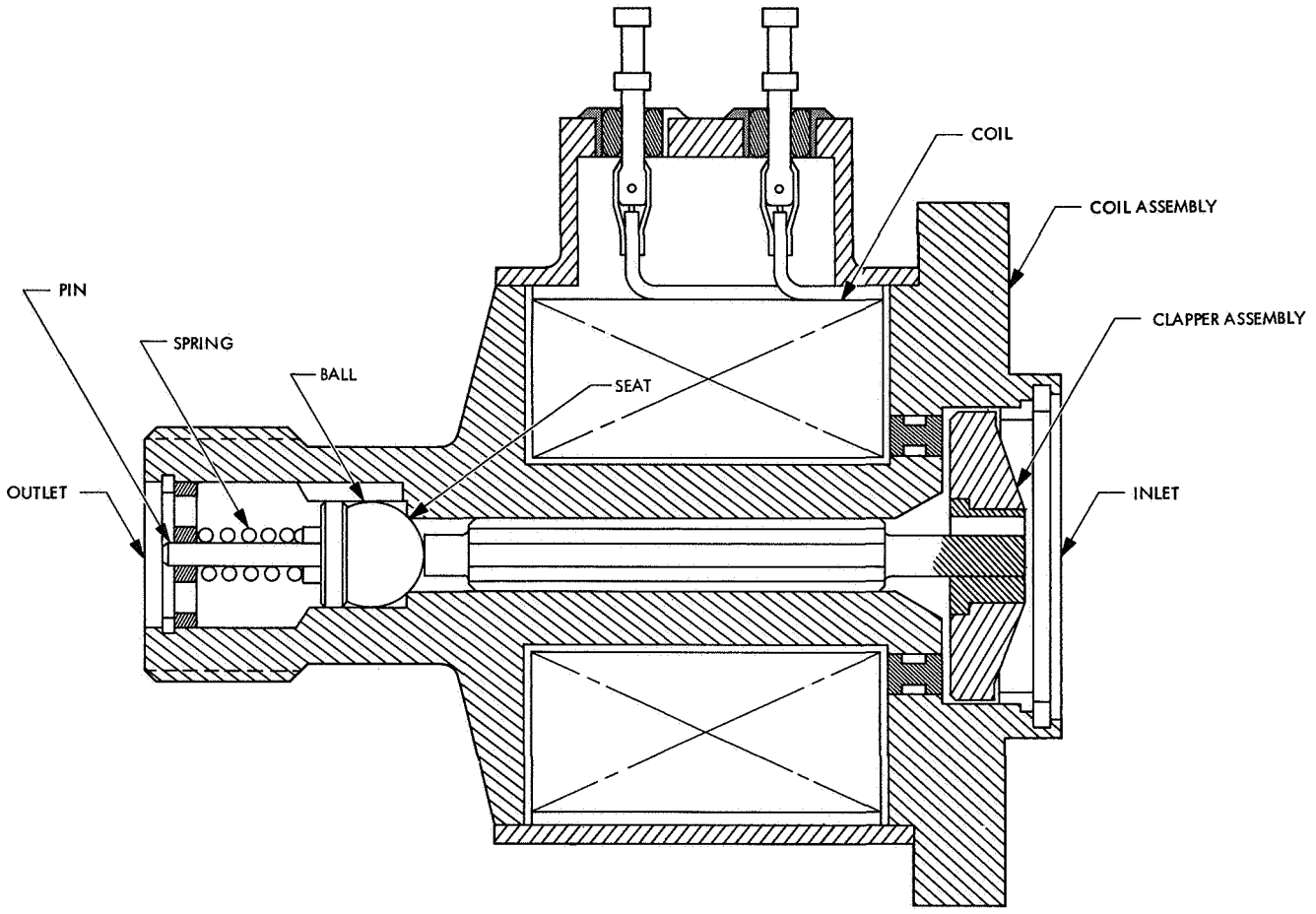


Fig. 3. Cutaway view of *Mariner IV* valve

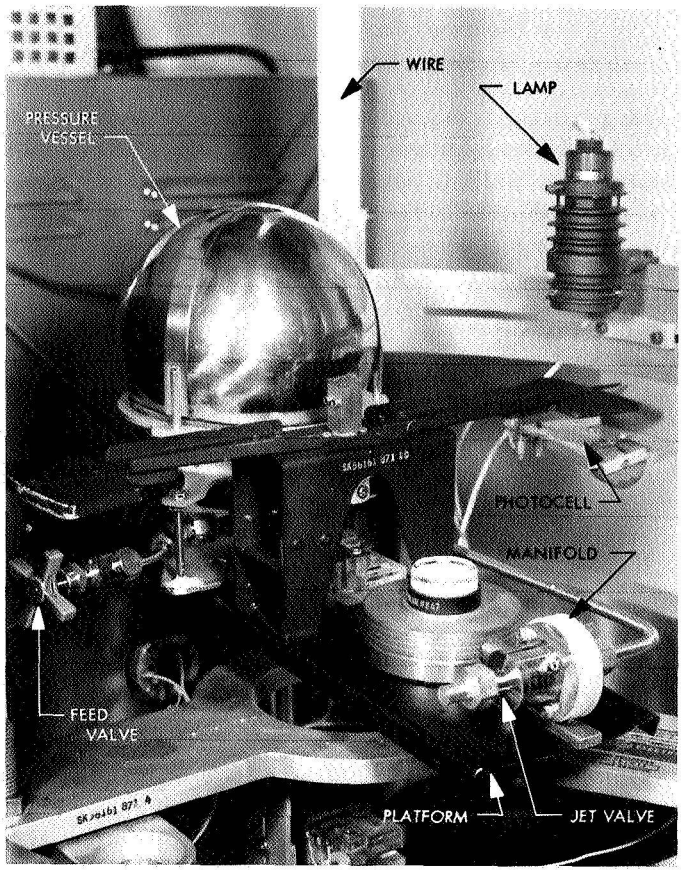


Fig. 4. Thrust platform and readout mechanism

The thrust platform and readout mechanism are shown in Fig. 4.

Three valves were tested at average leak rates of 3.8, 5.0, and 13.5 cm^3 (std)/h, respectively. The leak rate varied considerably during the test, especially for the latter valve. The results are shown in Fig. 5 along with Laumann's analytical results, showing considerable correlation. The results of this test program will enable the trajectory analysts to take into account realistic values of the thrust produced by leaking jet valves both in predicting the future spacecraft dynamic behavior and in analyzing the past spacecraft telemetered flight data.

C. Dynamic Simulation Utilizing Thrust Profile,

L. L. Schumacher

1. Introduction

Many new types of low-thrust reaction devices (thrusters) have been developed in recent years. In addition, many of the conventional thruster designs have been

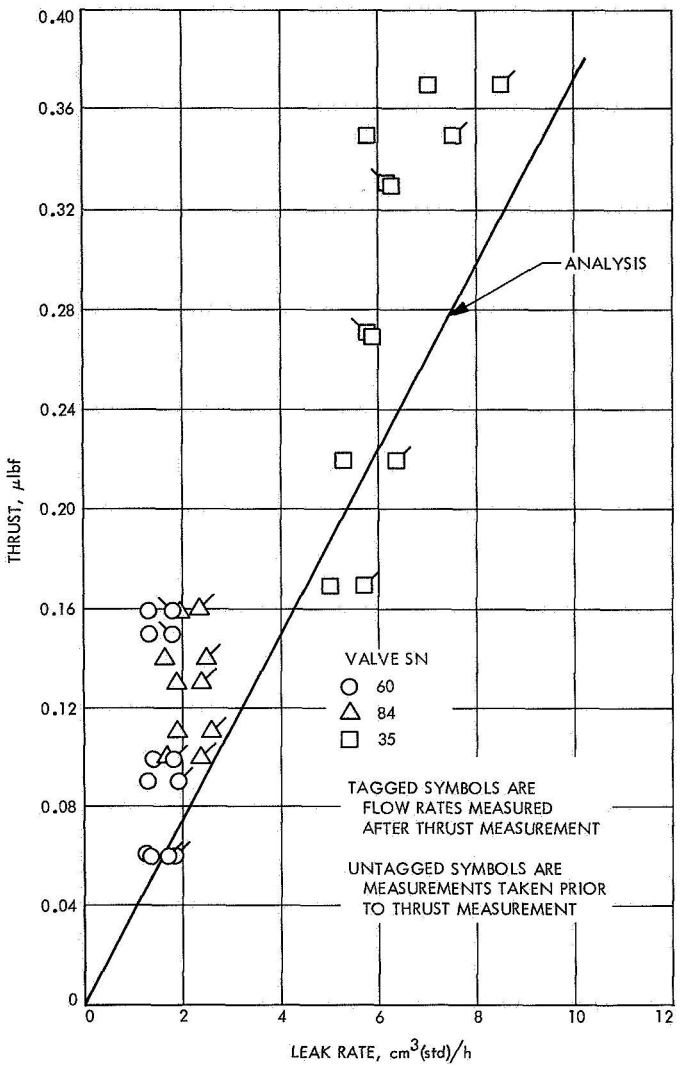


Fig. 5. Nozzle thrust versus leakage flow rate

modified using new thrust delivery techniques. In many cases, the transient thrust characteristics of these thrusters are radically different than the ideal or desirable thrust characteristics.

Initially, thruster selection for a given mission is made on the basis of the ideal or steady-state thrust requirements. If a thruster is considered which has transient thrust characteristics that deviate markedly from the ideal, it would be wise to examine the dynamic interaction of spacecraft, attitude-control systems, and thrusters.

To evaluate the effects which thruster transients have on a spacecraft, a DSL/90 digital simulation program was written. The program is capable of simulating attitude-control systems in a wide variety of spacecraft.

The following features are also realized in the simulation program:

(1) Rigid body model.

(2) Six degrees of freedom.

(3) Nonlinear sensor models.

(4) Three-axis thruster control system.

(5) Three-axis momentum wheel control system.

(6) Momentum wheel unloading system.

(7) Inter-axis coupling.

(8) Accurate model of thruster transient characteristics.

(9) Various attitude-control modes.

(10) Accurate gas consumption calculations in each mode.

This article elaborates on a few of the above features, and presents a sample of the results.

2. Thruster Transient Characteristics

In spacecraft dynamic analysis, the thruster transient of interest is the "thrust profile." This is the locus of thrust versus time. Figure 6 is a typical thrust profile of a reaction-jet type of thruster. The following are properties of interest in a thrust profile:

(1) T_1 is the time it takes to establish current in the inductive coil and overcome poppet inertia.

(2) T_2 is the time the thrust has reached its steady-state value.

(3) T_3 is the shutoff signal time. For various modes of operation the thruster systems are artificially rate compensated. This requires that a "minimum on time" should be greater than T_2.

(4) T_4 is the time when the valve has closed.

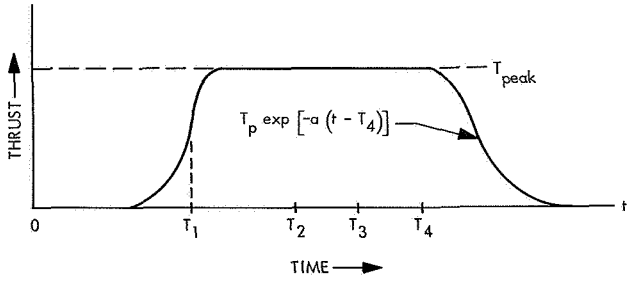


Fig. 6. Thrust profile

(5) $1/a$ is the time constant of an assumed exponential decay of thrust. This is called tailoff.

This program is designed so that the control torque profile is a very close approximation of the actual thrust profile times a constant (the moment arm). The above properties are only indicators to the experienced analyst that a potential stability problem exists. However, the danger of control instabilities as well as the possibility of introducing unacceptable pointing errors are eliminated using this program. (This, of course, depends on how closely the models approximate actual hardware.)

3. Attitude-Control Systems

The program has two basic control systems built in. Both the thruster control systems and the momentum wheel systems have the following error signal inputs available:

(1) Position sensor outputs.

(2) The output of rate-sensing devices such as gyros.

(3) Derived-rate type of compensation.

In addition, any combination of the above three can be obtained. In this manner, most conventional attitude-control modes can be simulated, and the control system parameters sized using linear analysis methods are subjected to the most realistic tests.

4. Results

The model chosen was the ATS-4 of Ref. 1. The control modes simulated were as follows:

(1) Initial rate reduction.

(2) Sun acquisition.

(3) Momentum wheel unloading during a commanded turn.

The thrust profile model is typical of a cold gas system, and similar to Fig. 6 (see Ref. 2).

In all modes the model was stable. However, Figs. 7 and 8 illustrate the effect of using a thrust profile as opposed to an ideal square-wave thrust form. Figure 7 plots principal axis rates versus time in the rate reduction mode. Indicated is the increased time necessary for the reduction in the pitch and roll axes using a model thrust profile other than the ideal. Figure 8 is a yaw plane during a sun acquisition. Here again, the model thrust profile, which is very near ideal, produces a phase plane

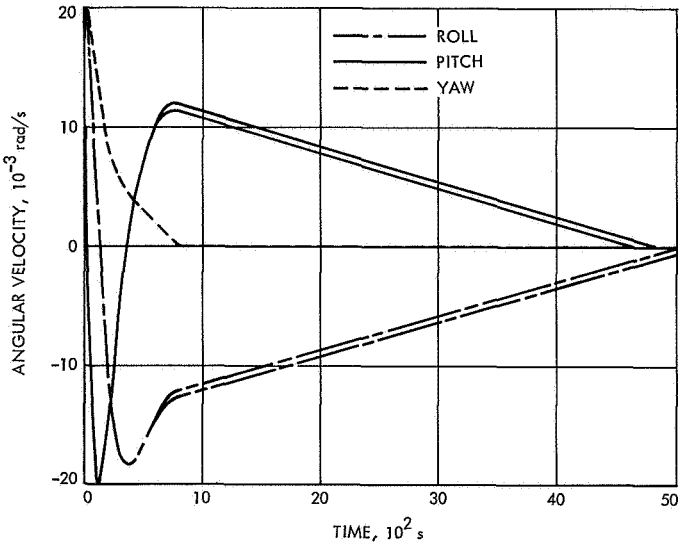


Fig. 7. Spacecraft rates

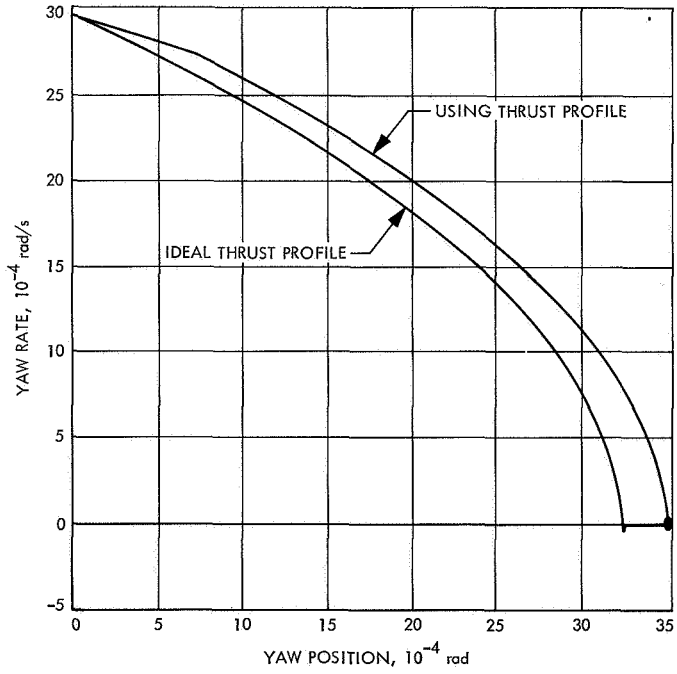


Fig. 8. Yaw phase plane

noticeably different than that produced by an ideal thrust form.

A more detailed treatment of this case can be found in a JPL internal document.[2]

[2]Schumacher, L. L., *Status of Computer Program for Thrust Profile Analysis*, May 8, 1969.

5. Conclusions

The computer program is functionally complete at this time. All modes of interest have been successfully simulated.

The major stumbling block is obtaining thrust profile information. When more information of this type is accumulated, significant improvement in the analysis of attitude-control systems using thrusters will be possible.

References

1. *An Advanced Study of Application Technology Satellite (ATS-4) Mission*, Vol. I of Books 1 and 2. General Electric Spacecraft Dept., Valley Forge Space Technology Center, Philadelphia, Pa., Nov. 9, 1966.
2. Turk, W., *Ranger Block III Attitude Control System*, Technical Report 32-663. Jet Propulsion Laboratory, Pasadena, Calif., Nov. 15, 1964.

D. Design and Prototype Development of the Pulse-Width Modulator and the Armature Current Direction Sensor for the Minimum-Energy Controller, *Y. E. Sahinkaya*

1. Introduction

This article summarizes the design calculations of the pulse-width modulator and the armature current direction sensor; laboratory test results on the integrated control system for the minimum-energy controller (SPS 37-56, Vol. III, pp. 138–144) are also presented.

2. Pulse-Width Modulator

The primary function of the pulse-width modulator is to supply firing signals to the silicon-controlled rectifier (SCR) gate circuits of the dc-to-dc voltage converters in such a way that the armature voltage of the motor is controlled according to the optimal control law during the motoring and generating modes of operation. The operation of the motor at any instant of time, in either the motoring or generating mode, depends on the difference between the instantaneous values of optimal control and back emf voltages. If the optimal control voltage is greater than the back emf voltage at any instant of time, the armature current direction sensor switches on the voltage supply for the control circuit of the first portion of the pulse-width modulator, which in turn yields the motoring mode of operation. The voltage supply for the control circuit of the second portion of the pulse-width modulator is switched off automatically. If the optimal control voltage at any instant of time is less than the back emf

voltage, the armature current direction sensor switches on the voltage supply for the control circuit of the second portion of the pulse-width modulator, which in turn yields the generating mode of operation. The voltage supply for the control circuit of the first portion of the pulse-width modulator is switched off automatically. The two portions of the pulse-width modulator have identical circuits.

The requirements to be satisfied by the pulse-width modulator are determined from the design considerations of the dc-to-dc voltage converters. The requirements are:

(1) Pulse frequency: 250 Hz.

(2) Minimum pulse-width: 400 μs.

(3) Maximum pulse-width: 3.6 ms.

(4) Maximum firing signal to SCR gate circuits: 6.8 V.

(5) Maximum current to SCR gate circuits: 40 mA.

(6) Nominal pulse-width of the firing signal to SCR gate circuits: 20 μs at 6.8 V.

The input voltage to the pulse-width modulator is related to the conduction duration of the main SCR by a linear relationship.

The values of the circuit elements of the pulse-width modulator are determined by utilizing the basic principles of linear circuit theory, solid-state device theory, and by carrying out the necessary experimentation on the prototype pulse-width modulator. Figure 9 shows the final values of the circuit elements of the pulse-width modulator.

In the design of the pulse-width modulator, the following procedure has been used:

(1) A unijunction relaxation oscillator (Ref. 1) is designed to generate a sawtooth waveform with a stable frequency of 250 Hz and a peak value of 6.8 V. This is accomplished by charging the 0.1-μF capacitor with a constant current from the collector circuit of M8. The emitter circuit of M5 provides the necessary bias level, which is linearly proportional to the optimal control voltage, for the sawtooth voltage waveform in order to obtain the pulse-width modulation action. Whenever the amplitude of the capacitor voltage exceeds the amplitude of the emitter bias voltage of M5 within the period of relaxation oscillation, the collector voltage of M7 is applied to the base circuit of M10 of the Schmitt trigger in such a way that the desired firing signal is generated across the primary winding of pulse transformer T1 (Fig. 9). The secondary winding of T1 is connected to the gate circuit of the main SCR of the dc-to-dc voltage step-down converter, i.e., motoring mode of operation. The circuit which includes all the elements between zener diode Z6 and pulse transformer T1 is the control circuit of the first portion of the pulse-width modulator.

(2) A Schmitt trigger, which is controlled by the Schmitt trigger of the control circuit, is designed to generate the desired firing signal to the primary winding of pulse transformer T2. The secondary winding of T2 is connected to the gate circuit of the commutating SCR of the dc-to-dc voltage step-down converter. The bias circuit of M12 is adjusted such that whenever the transistor M11 turns off, a firing signal is generated across the primary of T2. This in turn causes the commutating SCR to turn off the main SCR in the dc-to-dc voltage step-down converter.

The operation of the pulse-width modulator is described schematically in Fig. 10 for two particular idealized conditions. The conduction duration of the main SCR is linearly proportional to the average voltage available at the motor armature circuit.

3. Armature Current Direction Sensor

The armature current direction sensor controls the mode of operation for the pulse-width modulator. It consists of a comparator circuit and two Schmitt trigger circuits. The Schmitt triggers turn on or off the voltage supply to one of the control circuits of the pulse-width modulator depending on the polarity of the error signal between the sample signals of the optimal control and back emf voltages. The only requirement to be satisfied by the armature current direction sensor is the provision of a deadband between the positive and negative error signals. This is desired for reasons of reliability.

The values of the circuit elements of the armature current direction sensor are determined by utilizing the same principles as mentioned in *Subsection* 2. Figure 11 shows the circuit elements of the armature current direction sensor with their final values. The interface relationship between the configurations of the armature current direction sensor and both portions of the pulse-width modulator is also shown.

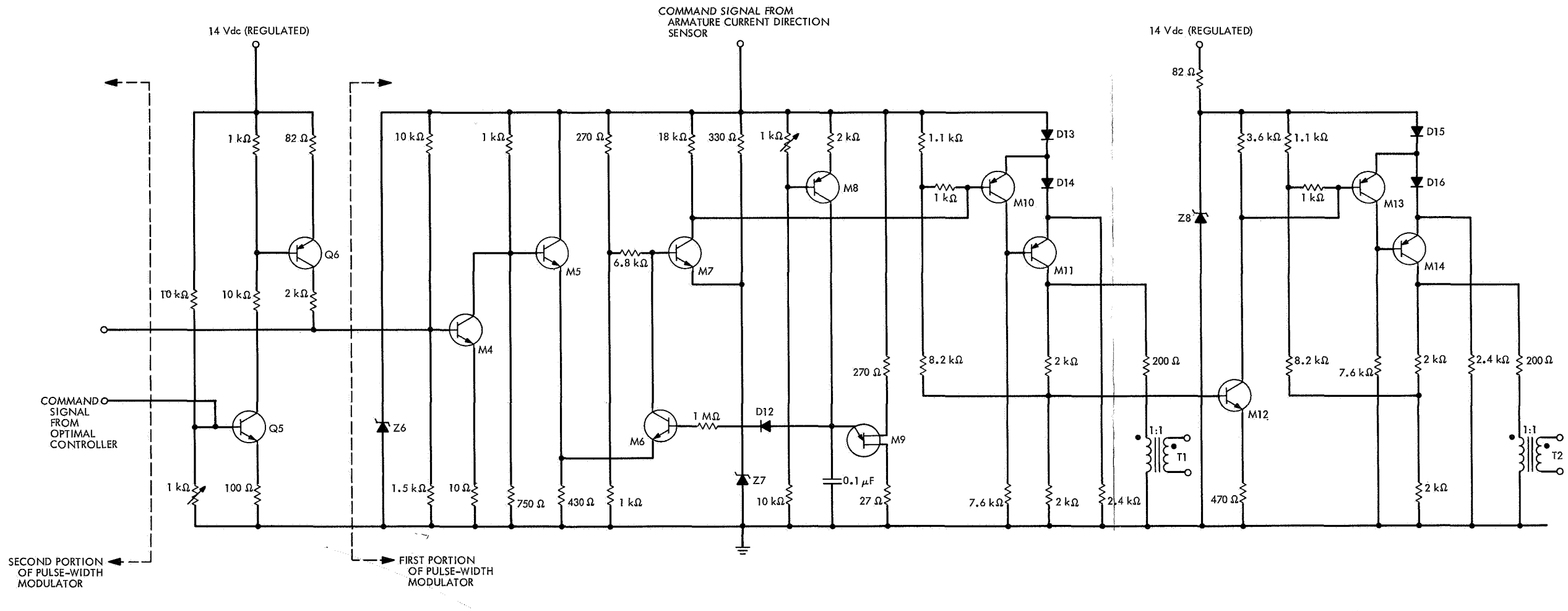


Fig. 9. Pulse-width modulator

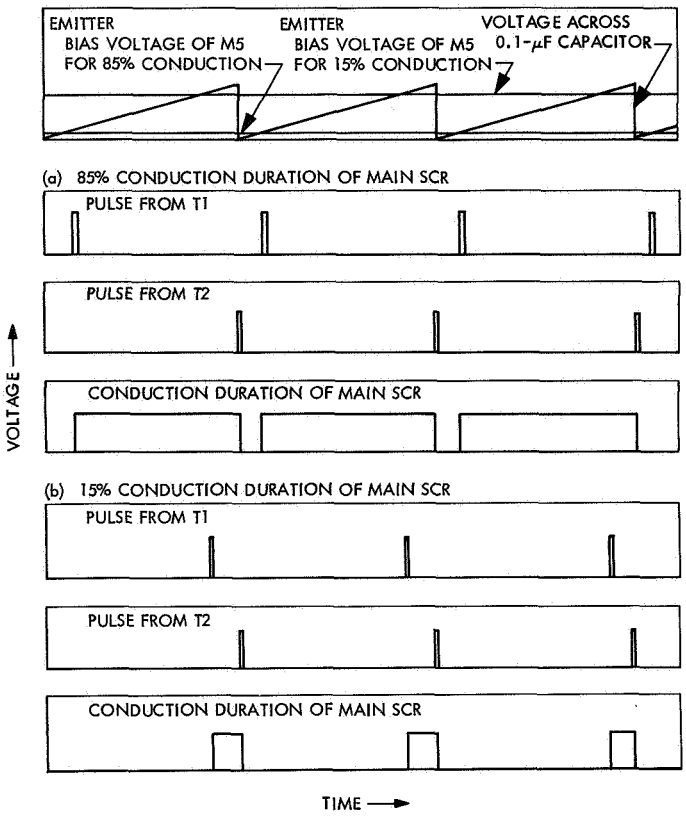


Fig. 10. Ideal characteristics of the pulse-width modulator

4. Laboratory Test Results

The prototype pulse-width modulator and armature current direction sensor has been tested under simulated laboratory conditions. It has been observed that all requirements of this subsystem as specified above are satisfied.

The salient features of the pulse-width modulator are studied by using a high-speed oscilloscope which is fitted with a camera. It has been determined that the pulse-width modulator is capable of varying the conduction duration of main SCRs from 10 to 90% of their maximum conduction durations during the motoring as well as generating modes of operation.

Figure 12 shows a typical sample of the firing characteristics of the pulse-width modulator for 20 and 80% conduction durations of the main SCR. Note that the negative values of the pulse voltages are caused by the free-wheeling action of the zener diodes which are placed in parallel with the secondary windings of T1 and T2. The relationship between the input voltage to the base circuit of Q5 and the conduction duration of the main SCR is shown to be approximately linear over the workable range.

5. Conclusions

The combined operation of the pulse-width modulator and the armature current direction sensor is vital to the operation of the minimum energy control system. Hence, every conceivable effort has been made to ensure that it will operate very satisfactorily at all times under expected working conditions. In this connection, the prototype model has been subjected to various tests in the last 4 to 5 weeks under laboratory conditions. The results, so far, have been very satisfactory.

Reference

1. *Transistor Manual, Unijunction Transistor Circuits*, Seventh Edition. General Electric Co., Syracuse, N. Y., 1964.

E. Digital Sun Sensor, *L. F. Schmidt*

1. Introduction

The sun sensors used on the *Mariner* series spacecraft are calibrated to provide accurate error angle outputs only near the deadband limits of ±0.25 deg. Spacecraft design studies for the outer-planet missions, including Grand Tour, indicate the need for sun sensors of a completely different type. This need is predicated by the requirement to operate over a very wide range of solar intensity and by the technique to be used for antenna pointing. Pointing of the spacecraft high-gain antenna is accomplished by biasing the output of the sun sensors in both pitch and yaw. The attitude-control system to null out this bias then turns the entire spacecraft to achieve antenna pointing. This technique requires high sun sensor accuracy over the total sun sensor field of view (FOV). Initial trajectory analysis data for the Grand Tour mission indicates that an FOV of ±3 deg in pitch and +24 and −12 deg in yaw will meet the mission requirements.

Two different types of optical systems for the sensor have been considered, one for each axis to minimize the size of the required size of the detector arrays. Design studies showed, however, that two very similar systems would result in acceptable detector sizes and minimize the volume of the sun sensor.

A single-axis breadboard having a ±3 deg FOV is being constructed. The information gained from testing will be applicable to the subsequent design of a two-axis engineering prototype.

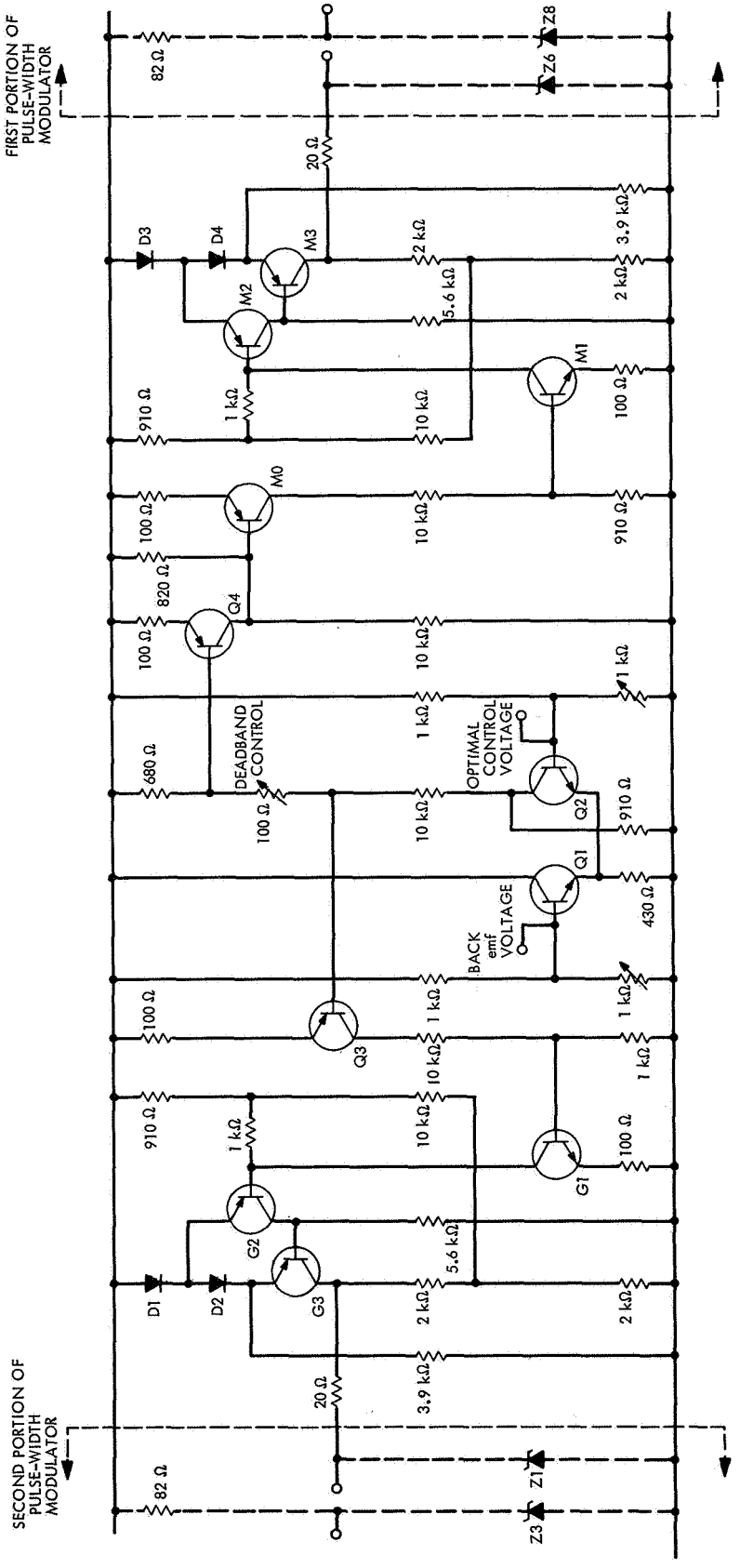


Fig. 11. Armature current direction sensor

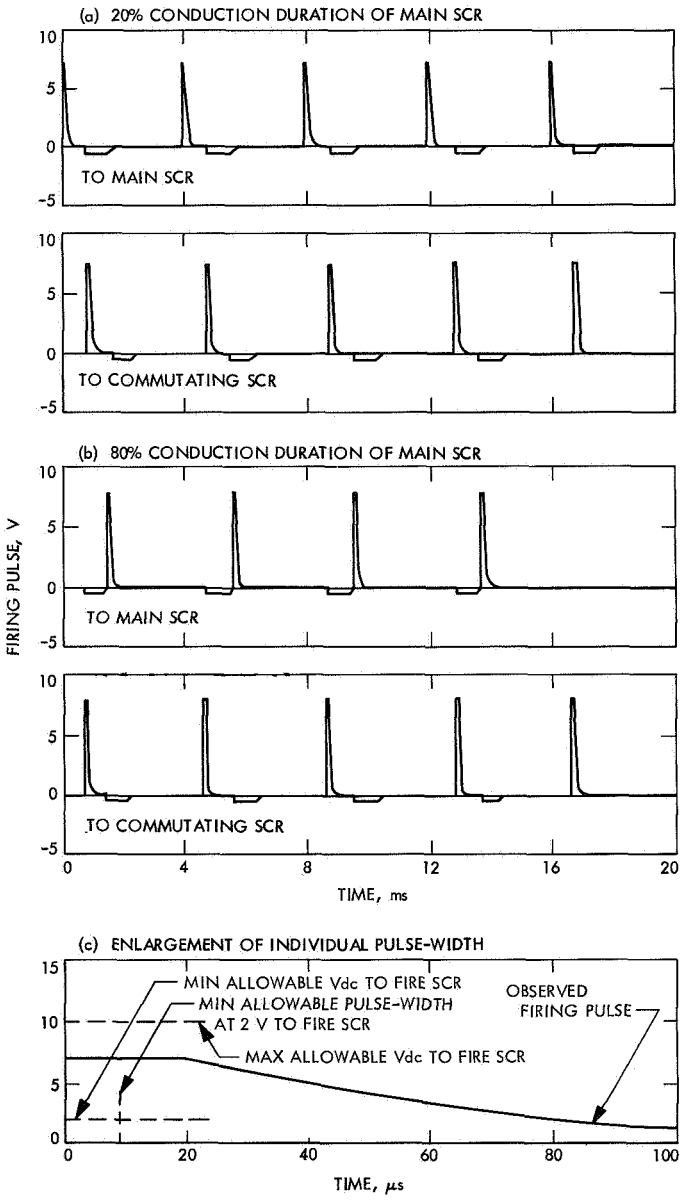


Fig. 12. Firing characteristics of the pulse-width modulator

2. Detector Design

A cadmium sulfide detector array has been designed. Cadmium sulfide detector material of the recrystallized type was chosen in preference to silicon because it has better resistance to radiation and high detectivity at low light levels.

The detector array is illustrated in Fig. 13. The illuminated strip is formed by imaging the sun on the detector using a cylindrical lens. As the spacecraft pitch attitude changes, the strip image of the sun is swept across the detector. This causes a negative or positive voltage to appear at each output terminal as controlled by the gray code pattern. The gray code is used instead of a binary code to prevent ambiguities that occur with binary coding because the code mask tolerancing and detector non-uniformities can result in simultaneous or reverse order switching of two or more bits. The gray code, sometimes referred to as reflected binary, requires only one bit to change state at one time.

As the illuminance decreases with increased distance from the sun, the illuminated portion of each cadmium sulfide strip increases in resistance. However, at the output electrodes, the voltage zero crossings shift only by the amount caused by the non-uniform characteristics of the detector strip pairs. Figure 14a illustrates a pair of detector strips required to produce one bit of the output word at 1 astronomical unit (AU) and the associated transfer function. Figure 14b illustrates the same information at 30 AU. The sun image width change causes large variations in the transfer function; however, the zero crossings occur at essentially the same position. The signal processing circuitry will generate a 1 or 0 output bit depending on whether the detector output is on the plus or minus side of zero voltage.

The detector design illustrated in Fig. 13 has been designed in detail and procurement of four detectors from a vendor is scheduled for July 1969.

3. Optical Design

A narrow slit aperture was considered in lieu of a cylindrical lens. The narrow slit approach would cause the detector to operate over an extreme variation in illuminance: 12,600 ft-cd at 1 AU and 14 ft-cd at 30 AU. The cylindrical lens approach reduces this incremental change at the detector to 12,600 ft-cd down to 840 ft-cd, as indicated in Fig. 14. The cylindrical lens also allows the range of illuminance to be shifted by means of aperture sizing. It allows a sealed design that is less susceptible to contamination and provides a surface for control of the spectral content of the energy entering the sensor, through means of an interference filter.

For the pitch axis, a cylindrical lens having a focal length of 3.5 in. was designed and evaluated using a JPL computer program.[3] It was found to have sufficient depth

[3]The description of this program is covered by Firnett, P. J., and Wilson, L. A., *FORTRAN Optical Lens Design Program*, Vols. I, II, and III, TR-67-700-10-2. Prepared under JPL contract by Informatics, Inc., Los Angeles, Calif., Oct. 1, 1968.

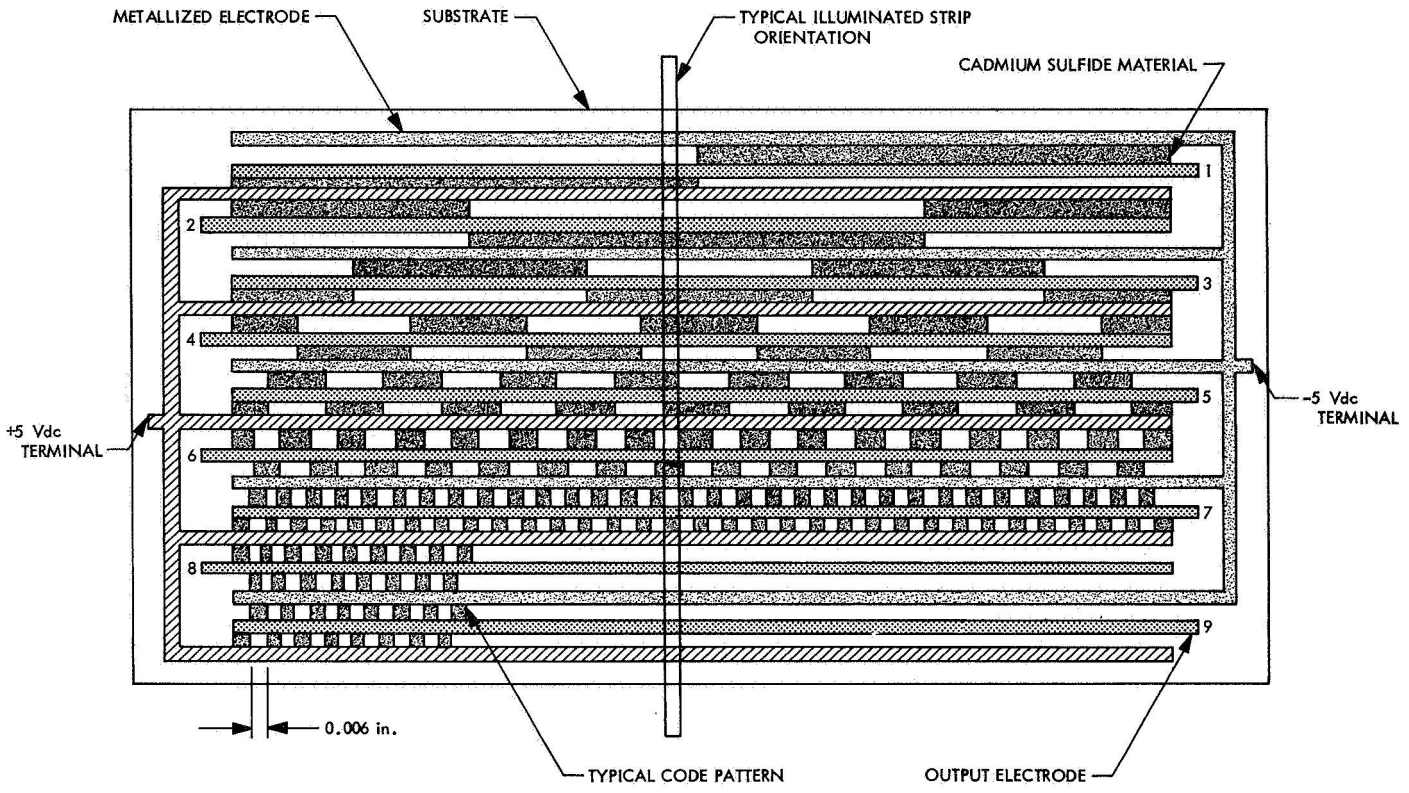


Fig. 13. Detector array

of field to allow use of a flat detector array at the image plane.

Use of the same techniques for the yaw axis would require a detector length of approximately 3.5 in. because of the increased FOV. The state of the art in recrystallized cadmium sulfide detectors allows fabrication of sizes up to 2 inches square. Uniform characteristics cannot be attained with larger detectors using current processing techniques.

For this reason, the optical configuration illustrated in Fig. 15 was evaluated. In this system, element 1 images the sun on a gray code mask deposited on the flat surface of element 2. Elements 2 and 3 image the front aperture on the detector surface with the energy passing through lens 4, which is an array of cylindrical lens elements. Each cylindrical element images a row of coded slots onto the appropriate cadmium sulfide strip. Two elements of the array are required for each output bit. Assuming a 10-bit word is required to cover the 39-deg total FOV, 20 cylindrical elements are required in the array. Design studies resulted in a required detector length in the pitch direction of approximately 3.2 in., even though the dimension along the yaw axis was reduced to 0.25 in. Although two separate detectors could be used, the sensor's physical size becomes unacceptably large and the optical system relatively complex when using this approach.

As was shown in Fig. 13, the gray code pattern is symmetrical in the direction of the cadmium sulfide strips. Using the optical technique illustrated in Fig. 16, a half-length pattern may be used. Lens A illuminates the detector when the solar input comes from the right of vector Z. Lens B illuminates the detector when the solar input comes from the left of vector Z. In order to generate the most significant bit, the associated cadmium sulfide strips are located near the edges of the substrate, one being energized by lens A and the other by lens B. This technique allows use of a detector of approximately 1.6 by 0.650 in. The length of the sensor along the optical path can also be reduced to about one-half compared to the aforementioned system.

4. Electronic Circuitry

The signal processing circuitry is described by the following conceptual philosophy as related to the attitude-control constraints. In general terms, the gray code word

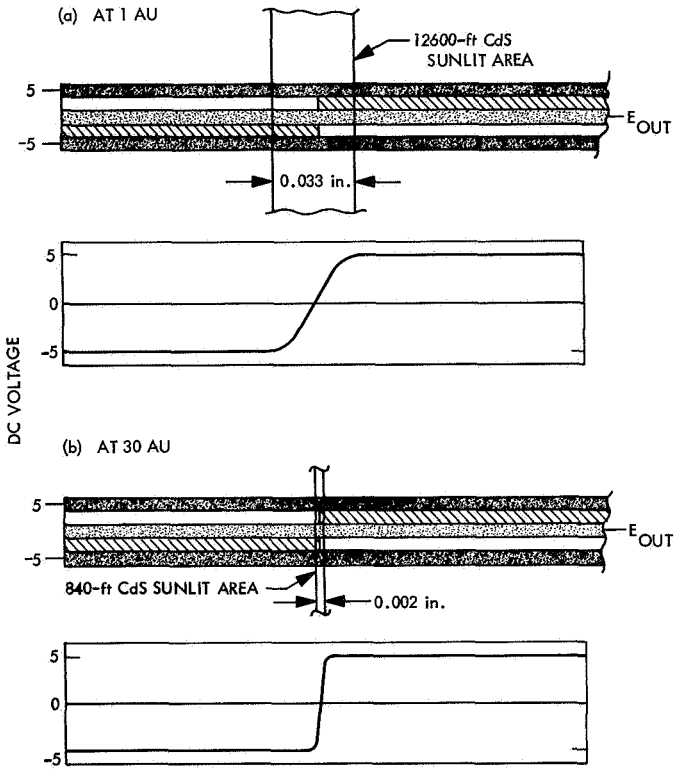


Fig. 14. Output function and sun image width

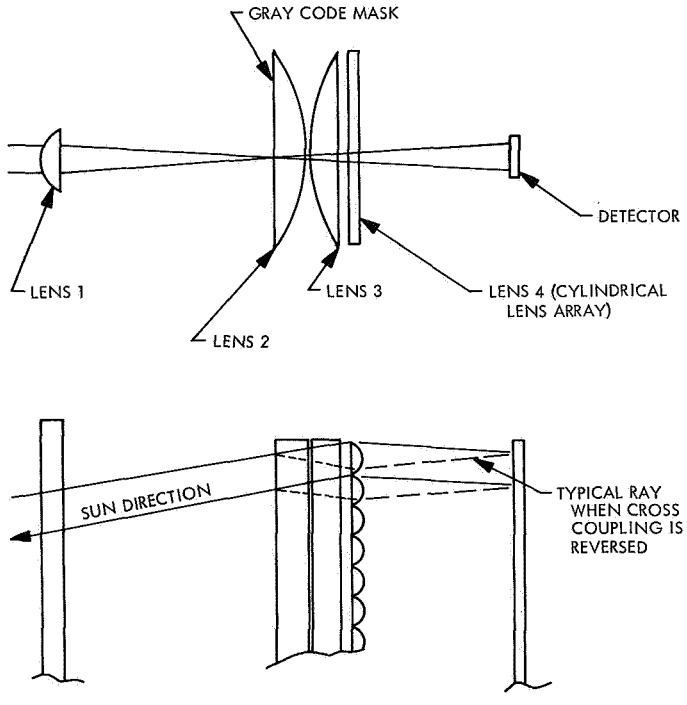


Fig. 15. Optical system yaw axis field lens technique

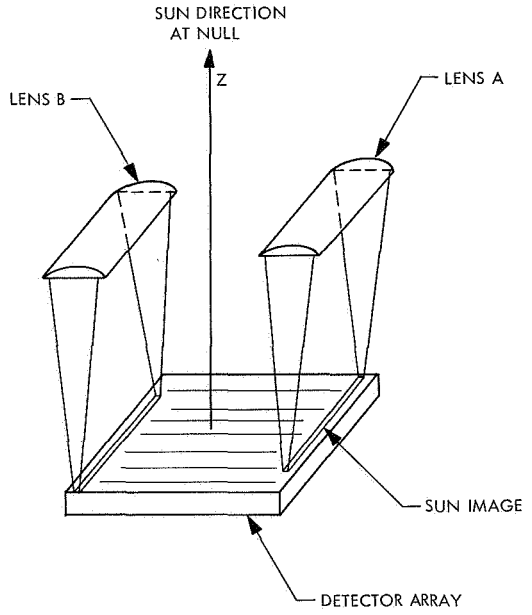


Fig. 16. Optical system yaw axis dual objective lens technique

from the detector is fed into a storage register in a parallel manner. Each bit is transmitted into a shift register, also in parallel. The shift register can be read out by the attitude-control logic, while it maintains a disable command to prevent the word from changing. The readout is accomplished serially to reduce the number of wires between the sun sensor and the attitude-control logic. The wire cable will be on the order of 20 ft in length. The serial readout reduces the estimated number of wires from 24 to 7.

5. Breadboard Fabrication and Testing

The construction of the optical mechanical breadboard is 95% complete with the exception of the detector; breadboarding of the electronics has started.

Initial testing of the breadboard is scheduled to begin in July upon receipt of the detectors.

The first test module to be constructed will be designed to monitor the output at the detector. Testing of this partial system is planned to proceed in parallel with the development and testing of the complete test console.

6. Accuracy

The estimated resolution and accuracy are shown in Fig. 17. At 1 AU, the resolution is limited to 0.2 deg because the sun's angular diameter produces an illuminated strip too wide to operate the detector's three least-

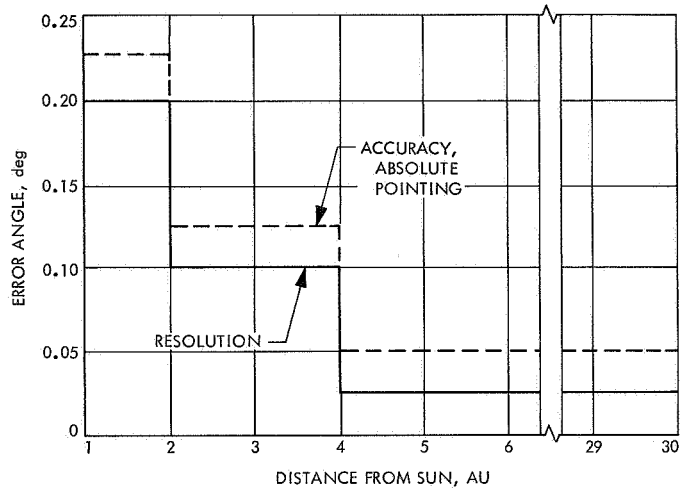


Fig. 17. Estimated accuracy/distance from sun

significant bits. At approximately 2 AU, two significant bits are inoperative for the same reason. At approximately 4 AU, the two remaining significant bits operate correctly since the code slots that generate these bits are equal in width.

In order to achieve the 0.025-deg resolution bit at 4 AU, two pairs of cadmium sulfide strips are employed (output channels 8 and 9 in Fig. 13). The coded pattern of one leads the other by one-half of the code slot width. These two detector outputs 8 and 9 are combined with the appropriate digital logic elements to produce 1 output bit. The complete error angle is then expressed by an 8-bit word.

The absolute accuracy is dependent on the resolution, the mechanical tolerances to which the code mask can be manufactured, the variation in detector characteristics, variation in signal processing circuitry characteristics, stray light effects, the mechanical stability of the optical alignment, and cross-coupling effects. Tests of the breadboard sensor will provide an empirical evaluation of the relative importance of these error sources.

F. On-Line Estimation of Atmospheric Parameters by an Approximate Nonlinear Filtering Technique, *A. K. Bejczy*[4]

1. Introduction

Plans for landing missions to distant atmospheric planets (e.g., to Mars) generated interest in analytical studies on optimal or near-optimal self-contained terminal guidance and control systems that can operate in an unmanned landing vehicle. Because of the imperfectly known and unpredictable planetary atmospheric conditions, an important building block of such a guidance and control system is a filtering scheme by which the dynamically relevant atmospheric parameters can be estimated on-line (i.e., in real time), based partly on the current on-board measurements of the atmospheric flight data and partly on the *a priori* knowledge of the physics of the atmospheric flight. There is, however, a severe practical constraint: the design of an appropriate filter scheme must not be overly complex to cope with the memory and computing capacity of a weight-limited on-board flight computer.

On-line estimation of the dynamically significant atmospheric parameters is essentially a nonlinear filtering problem because of the nonlinear mathematics associated with the dynamical and measurement equations of the atmospheric flight of a vehicle. In a previous communication,[5] a successful approximate nonlinear filtering technique was developed for sequential (i.e., on-line) estimation of state variables of a ballistic vehicle operating in an imperfectly known atmosphere. (The developed technique resulted in 90–95% reduction of mathematical operations that were needed to mechanize a corresponding nonlinear full filter scheme. Moreover, the approximate nonlinear filter provided better tracking performance than the corresponding full filter.) The developed approximation technique was restricted, however, to estimation problems with *all* state variables available for measurements. Clearly, this is not the case when physical constants (e.g., atmospheric density and density profile) must be estimated on-line by a dynamical filtering technique. Thus, to cope with such types of filtering problems when *not all* states are available for measurements, the approximate nonlinear filtering technique developed in Footnote 5 must suitably be reformulated and extended, if possible.

In this article, a suitable reformulation and extension of the approximate nonlinear filtering technique of Footnote 5 is discussed and demonstrated. Aided by extensive digital simulation studies, it has been found that a properly constructed constant-gain (or constant "covariance") nonlinear filter scheme can also be successfully applied to on-line estimation of atmospheric parameters. Because of

[4]Research completed during the tenure of a National Research Council Postdoctoral Resident Research Associateship at JPL, supported by NASA.

[5]Bejczy, A. K., *Sequential Estimation of States of a Ballistic Vehicle in Imperfectly Known Atmosphere*, Technical Report 32-1405. Jet Propulsion Laboratory, Pasadena, Calif. (in press).

its structural simplicity, this filter scheme can be implemented without requiring excessive computing capacity on board a landing vehicle. Again, the constant-gain (approximate) nonlinear filter scheme provides better tracking of the atmospheric parameters than the corresponding nonlinear full filter. For the sake of brevity, only a restricted vertical descent problem is presented.

2. Vertical Descent Through an Imperfectly Known Atmosphere

The atmospheric drag parameter is by definition a constant but due to its imperfectly known value it will be defined as a dependent state variable having zero time derivative. Adopting the quadratic drag law and the inverse-exponential atmospheric density distribution, the vertical motion of the center of gravity of a vehicle in free-fall is described by the following system of ordinary nonlinear differential equations:

$$\dot{x}_1 = -x_2 \tag{1}$$

$$\dot{x}_2 = g - kx_3x_2^2 \exp(-bx_1) + n \tag{2}$$

$$\dot{x}_3 = 0 \tag{3}$$

where the dot denotes time derivative, and

x_1 = altitude

x_2 = velocity

x_3 = ground-level atmospheric density

g = acceleration of gravity, a constant

b = scale factor, a constant

k = a constant, related to mass and geometrical design of vehicle; hence, $K = kx_3$ = atmospheric drag parameter

n = unknown dynamic inputs, inclusive stochastic forces

Thus, in this restricted formulation, the estimation of the imperfectly known atmospheric drag parameter is equivalent to estimating the ground-level density of the atmosphere during the vehicle's atmospheric descent.

The observation vector is assumed to be linearly related to the states

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \text{noise} \tag{4}$$

where y_1 and y_2 can be outputs of a radar altimeter and radar velocity sensor, respectively.

Now, the problem is to estimate the current "best" value of the state vector $x = \{x_1, x_2, x_3\}$ in a sequential manner, given noisy linear observations on x_1 and x_2. Note that there are no observations on x_3.

3. Maximum Principle Least-Squares Nonlinear Filter

In the case of a linear observation vector h, the continuous formulation of the multidimensional maximum principle least-squares (MPLS) nonlinear filter (Ref. 1) has the structure

$$\dot{z} = f(z) + 2P\bar{h}_z Q\,[y(T) - h] \tag{5}$$

$$\dot{P} = f_z P + P\bar{f}_z - 2P\bar{h}_z Q h_z P + R \tag{6}$$

where

z = estimated value of n-dimensional state vector x

T = running estimation (and observation) time

$f(z)$ = system equations, n-vector function

$h(z)$ = Hz, observation equations, m-vector function, $m \leq n$, and H is an $m \cdot n$ matrix

f_z, h_z = Jacobians of f and h, respectively; the bar denotes their transposes

Q = quasi-norm factor, positive semi-definite $m \cdot m$ matrix related to the observations; we take $Q = I$, identity matrix

R = $(\frac{1}{2})\,W^{-1}$, inverse of positive definite $n \cdot n$ quasi-norm matrix W related to system equations

P = gain (or approximate "covariance"), $n \cdot n$ symmetric matrix

Thus, the full MPLS nonlinear filter scheme for estimating altitude, velocity, and ground-level density, given noisy altitude and velocity measurements, is expressed by the following system of ordinary nonlinear differential equations:

$$\dot{z}_1 = -z_2 + 2P_{11}(y_1 - z_1) + 2P_{12}(y_2 - z_2) \tag{7}$$

$$\dot{z}_2 = g - z_3 z_2^2 A + 2P_{12}(y_1 - z_1) + 2P_{22}(y_2 - z_2) \tag{8}$$

$$\dot{z}_3 = 2P_{13}(y_1 - z_1) + 2P_{23}(y_2 - z_2) \tag{9}$$

$$\dot{P}_{11} = -2P_{12} - 2P_{11}^2 - 2P_{12}^2 + R_{11} \tag{10}$$

$$\dot{P}_{12} = -P_{22} + bAz_3z_2^2P_{11} - 2Az_3z_2P_{12} - Az_2^2P_{13} - 2P_{11}P_{12} - 2P_{12}P_{22} + R_{12} \quad (11)$$

$$\dot{P}_{13} = -P_{23} - 2P_{11}P_{13} - 2P_{12}P_{23} + R_{13} \quad (12)$$

$$\dot{P}_{22} = 2bAz_3z_2^2P_{12} - 4Az_3z_2P_{22} - 2Az_2^2P_{23} - 2P_{12}^2 - 2P_{22}^2 + R_{22} \quad (13)$$

$$\dot{P}_{23} = bAz_3z_2^2P_{13} - 2Az_3z_2P_{23} - Az_2^2P_{33} - 2P_{12}P_{13} - 2P_{22}P_{23} + R_{23} \quad (14)$$

$$\dot{P}_{33} = -2P_{13}^2 - 2P_{23}^2 + R_{33} \quad (15)$$

where $A = k\exp(-bz_1)$.

It must be observed that Eqs. (7–15) form a coupled system of ordinary nonlinear differential equations of *stochastic* character; the "state estimator equations" (Eqs. 7–9) are of stochastic nature because of the noisy y_i forcing terms, while the gain (or "covariance") equations (Eqs. 10–15) are of stochastic character because of the explicit appearance of z_i as coefficients for some of the P_{ij} terms.

The characteristic feature of Eqs. (7–9) is their asymmetry in the P_{ij} terms: there is no P_{13} term in Eq. (7), no P_{23} term in Eq. (8), and no P_{33} term in Eq. (9). Thus, the "diagonal reduction technique" developed previously (Footnote 5) can not be applied immediately for these filter equations. The only immediate simplification that might be introduced into Eqs. (7–15) is to put $\dot{P}_{12}=P_{12}=0$. This simplification can be motivated by the "double role" of the P_{12} term in Eqs. (7) and (8): P_{12} transmits contributions from velocity measurements to estimating altitude, as well as transmitting contributions from position measurements to estimating velocity.

Extensive digital simulation studies were carried out on Eqs. (7–15) by solving them for different combinations of $z_i(0)$, $P_{ij}(0)$, and R_{ij} values applying gaussian and/or uniform noise distribution. Regarding the filter's response characteristics, the simulation studies clearly displayed several important features of Eqs. (7–15).

(1) The filter's tracking performance remains unaltered or, in some cases, it even improves when the simplification $\dot{P}_{12} = P_{12} = 0$ is introduced into the filter scheme.

(2) Some of the P_{ij} terms settle down on constant values very soon after the initiation of the filtering process, while the other P_{ij} terms vary very little with time. Thus, after a short transient time, the gain equations exhibit an approximately algebraic structure. Note that this algebraic structure is nonlinear since the $\dot{P}_{ij} = 0$ equations remain second order in the P_{ij} terms. Note also that these nonlinear algebraic equations have time-varying coefficients because of the time-varying character of z_i.

(3) The transients of the gain equations cope very closely with the transients of the estimated trajectories on the time scale; i.e., the filter tracks the "true" trajectories beginning from that time point where the gains settle down on constant values or where the gains start to vary very little.

(4) The settled-down values of the P_{ij} terms strongly depend on the selected R_{ij} values. As a rule, higher R_{ij} values yield higher steady-state values for P_{ij}. It is also noted that high R_{ij} values result in shorter transients than the low ones, while the low R_{ij} values yield smoother estimates (i.e., less error) than the high ones.

(5) The filter's response is relatively insensitive to the initial error distribution.

It must be noted, finally, that all these characteristics of Eqs. (7–15) closely parallel the results obtained earlier for such cases where the state estimator equations were symmetric in the P_{ij} terms.

4. Constant-Gain Nonlinear Filter

Motivated by the response characteristics of Eqs. (7–15), a constant-gain diagonal filter was constructed by putting $\dot{P}_{11} = \dot{P}_{12} = \dot{P}_{13} = \dot{P}_{22} = \dot{P}_{23} = \dot{P}_{33} = 0$ in the filter equations and applying the settled-down values and/or the average steady-state values of the P_{ij} terms as constants in the state estimator equations from the initiation of the filtering process. Thus, the on-line solution (or mechanization) of all gain equations is omitted in this approximate filter scheme. The constant-gain filter for sequentially estimating altitude, velocity, and ground-level atmospheric density then reads

$$\dot{z}_1 = -z_2 + 2P_{11}^*(y_1 - z_1) + 2P_{12}^*(y_2 - z_2) \quad (16)$$

$$\dot{z}_2 = g - Az_3z_2^2 + 2P_{12}^*(y_1 - z_1) + 2P_{22}^*(y_2 - z_2) \quad (17)$$

$$\dot{z}_3 = 2P_{13}^*(y_1 - z_1) + 2P_{23}^*(y_2 - z_2) \quad (18)$$

where now the P_{ij}^* terms are constants.

A further simplification can be obtained in the constant-gain nonlinear filter scheme by putting $P^*_{12} = 0$ in Eqs. (16–18).

Digital simulation studies were carried out on the constant-gain nonlinear filter by solving Eqs. (16–18) for different sets of values for the P^*_{ij} constants and by applying gaussian and/or uniform noise distribution for the dynamic noise as well as for the observation noise. The main results of the digital simulation studies can be summarized as follows:

(1) The constant-gain nonlinear filter has better response characteristics than the corresponding nonlinear filter with time-varying gains.

(2) The filter's tracking performance is determined by the selected values for the P^*_{ij} constants: high P^*_{ij} values shorten the transient time for the estimated trajectories and increase the amplitude of the transients; low P^*_{ij} values yield smoother estimates, i.e., they yield less steady-state error after the transient part has been passed.

(3) There is a certain range for the individual P^*_{ij} values outside of which the filtering process loses its meaning: the estimated trajectories may diverge, the estimation errors may no longer be distinguishable from the measurement errors (MEs), or the estimation errors may exceed the measurement errors. The range for the individual P^*_{ij} values, within which the constant-gain nonlinear filter yields converging and sufficiently smooth estimates, is typically of one or two orders of magnitude.

(4) The initial error distribution, i.e., the assumed value of $z(0)$, affects only the transients of the estimated trajectories: $z(0) = 0$ yields the longest transient time, while any other physically meaningful $z(0)$ tends to shorten it.

Some representative results regarding the tracking performance of the constant-gain nonlinear filter are depicted in Figs. 18 and 19.[6] Figures 18a, b, and c refer to a filter scheme with $P^*_{12} = 0$, while Figs. 18d, e, and f depict the estimated trajectories belonging to the same cases but with $P^*_{12} \neq 0$. It can be seen that the only difference between the performances of the two filter schemes is in the transients of the estimated trajectories: the simplest filter scheme (where $P^*_{12} = 0$) tends to yield shorter transient time and smaller transient amplitudes for the estimated trajectories than the slightly more complex filter scheme (where $P^*_{12} \neq 0$). It is clear from these figures, however, that after 8–15 s of transient time both filter schemes yield stable and sufficiently smooth estimates of the "true" value of the ground-level atmospheric density, which was considered as unknown or known only with a high degree of uncertainty at the initiation of the filtering process. The 8–15 s transient time means that the filter must be in on-line operation in 8–15 s of real flight time to obtain a steady-state estimate on the ground-level atmospheric density. It must be observed that the filter tracks the "true" values of altitude and velocity much faster (typically, there is only 1 to 2 s of transient time for the estimated altitude and velocity) than the "true" value of the ground-level atmospheric density. The difference in the transient time is because the altitude and velocity are measured states, while the ground-level atmospheric density is not available for measurement; its value is estimated through altitude and velocity measurements.

Figure 19a shows how the "time histories" of the estimates on the ground-level atmospheric density are affected by changing the values of the constant gains in the filter scheme: the transient time can be decreased by increasing the relevant constant gains (i.e., P^*_{13} and P^*_{23}) but at the expense of increased steady-state estimation error (SSEE). Note, however, that the steady-state estimation error is also strongly dependent on the noise level of the altitude and velocity measurements. Thus, decreased noise level in the altitude and velocity measurements would yield, for higher constant filter gains, shorter transient time and less steady-state estimation error. For instance, 0.5% error in altitude and velocity measurements would yield, for the same constant gains as in Fig. 19a, only 5 to 6 s of transient time and only 3 to 4% of steady-state estimation error for the value of the ground-level atmospheric density.

Figure 19b gives the tracking reliability of the constant-gain nonlinear filter. With the same set of properly selected constant gains, the filter yields the same excellent tracking of the "true" values of the ground-level atmospheric density in a reasonably wide numerical range of the "true" parameter value.

[6] In the simulation studies the terminal phase of a typical Mars soft-landing problem was considered: a 1000-lb lander at 30,000-ft altitude with 800–900-ft/s velocity; the ground-level atmospheric density was assumed to be somewhere in the range of 0.5×10^{-5} to 3.5×10^{-5} g/cm^3; for the inverse-exponential atmospheric density distribution scale factor, a value of 2.15×10^{-5} ft^{-1} was used.

TRAJECTORY	$z_1(0)$, 10^3 ft	$z_2(0)$, 10^2 ft/s	$z_3(0)$, 10^{-5} g/cm^3
①	25	8	3.5
②	35	10	3.5
③	25	10	0.5
④	35	8	0.5
⑤	0	0	0

$P_{11}^* = 7.0$ $P_{22}^* = 9.0$ $P_{13}^* = 8 \times 10^{-8}$ $P_{23}^* = -4 \times 10^{-6}$

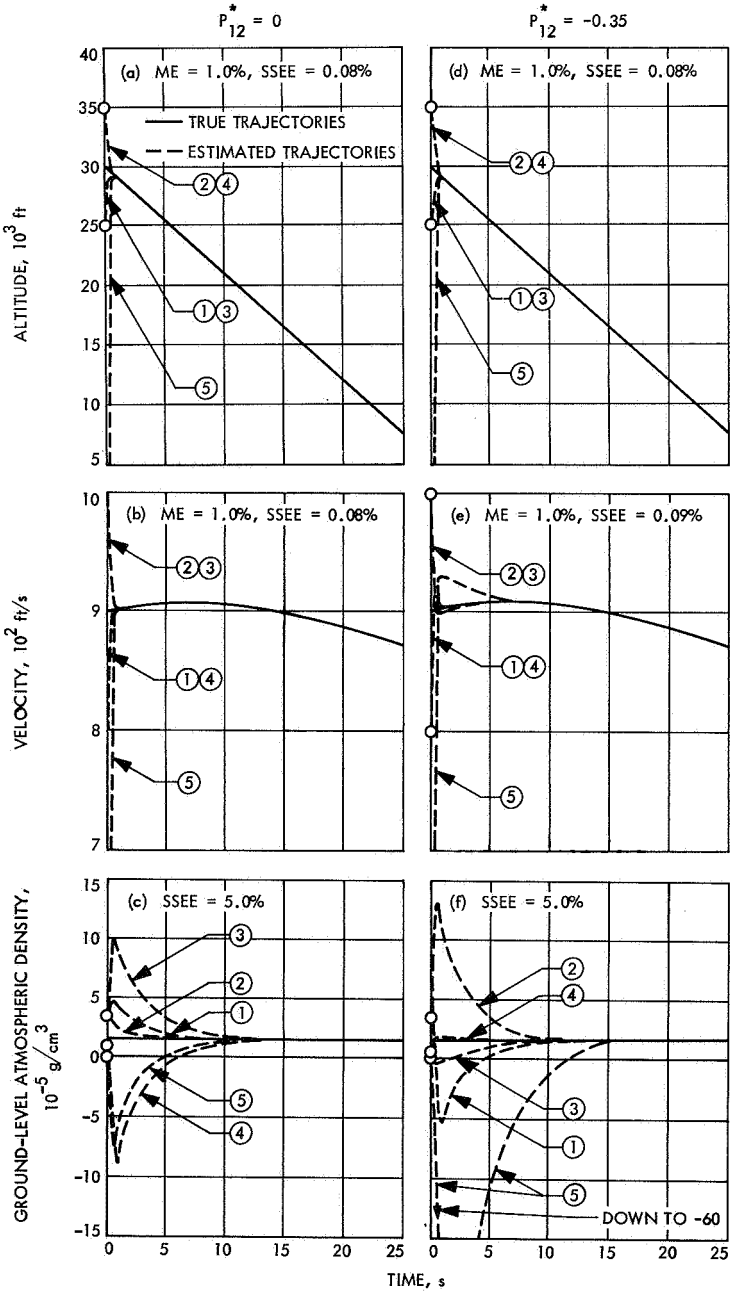


Fig. 18. Time histories of states and one parameter estimated on-line by a constant-gain nonlinear filter

TRAJECTORY	P^*_{11}	P^*_{22}	P^*_{13}	P^*_{23}
Ⓐ	7.0	9.0	8×10^{-8}	-4×10^{-6}
Ⓑ	7.0	10.0	5×10^{-7}	-1×10^{-5}
$z_1(0) = z_2(0) = z_3(0) = 0$				

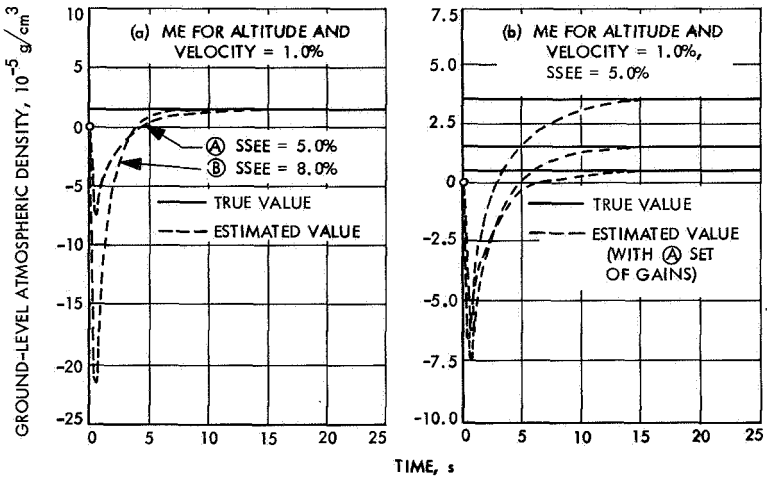


Fig. 19. Time histories of on-line estimated atmospheric parameter for different constant nonlinear filter gains and different "true" atmospheric parameter values

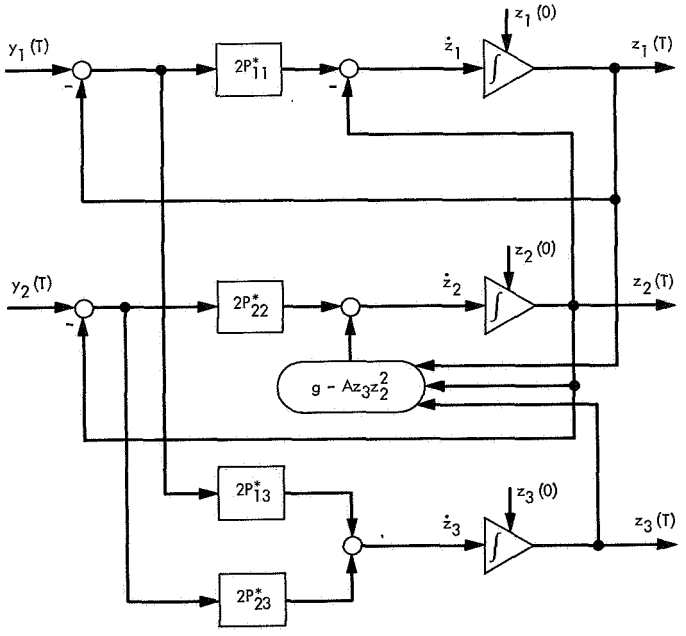


Fig. 20. Functional configuration of constant-gain nonlinear filter scheme

Figure 20 displays the functional configuration of the constant-gain nonlinear filter scheme given by Eqs. (16–18) for sequentially estimating altitude, velocity, and ground-level atmospheric density during vertical descent of a soft-landing given noisy altitude and velocity measurements. Clearly, the functional structure of this filter scheme is much simpler than the functional structure of the corresponding nonlinear filter with time-varying gains. The mechanization of the constant-gain nonlinear filter can be accomplished by 90% reduction of the mathematical operations that were needed to mechanize the corresponding nonlinear filter with time-varying gains. (It must be noted that, for the cases depicted in Figs. 18 and 19, the simulation time of the constant-gain nonlinear filter on an IBM 7094 computer is shorter than the actual simulated flight time!)

In conclusion, it is emphasized that, at the present stage of development, there is no way other than simulation studies for analyzing the performance characteristics of the constant-gain (or of any approximate) nonlinear filter. Thus, the approximations developed and discussed should primarily be viewed in relation to the particular application treated in this article. This caution is generally justified when the problem is nonlinear since, in such cases, the application itself also enters into the analysis and design of the filter scheme (Refs. 2 and 3).

References

1. Detchmendy, D. M., and Sridhar, R., "Sequential Estimation of States and Parameters in Noisy Nonlinear Dynamical Systems," *J. Basic Eng.*, Vol. 88, Series D, No. 2, pp. 362–369, 1966.
2. Kushner, H. J., "Approximations to Optimal Nonlinear Filters," *IEEE Trans. Automat. Contr.*, Vol. AC-12, pp. 546–556, 1967.
3. Schwartz, L., and Stear, E. B., "A Computational Comparison of Several Nonlinear Filters," *IEEE Trans. Automat. Contr.*, Vol. AC-13, pp. 83–86, 1968.

G. Solar Electric Spacecraft Attitude Control With Flexible Solar Arrays, *G. E. Fleischer*

1. Introduction

A "hybrid coordinate" approach to the dynamic modeling of spacecraft structures has been developed by Likins and Gale (Ref. 1). The particular advantages of this method include the capability of handling nonlinear mechanisms such as momentum wheels and/or the relatively large motions (translational or rotational) of *rigid* appendages along with the relatively small deformations of *flexible* appendages. The only requirement for applying the hybrid coordinate approach is that the rigid base, to which the flexible (and rigid) appendages are attached, undergo *small* motions or have nearly constant angular velocity. Normal modes of vibration can then be found to describe the flexible body motions of the system.

A hybrid coordinate model of a solar electric spacecraft configuration was derived and programmed for computer solution. This article describes some of the results.

2. Computer Simulations

The dynamic model described here was applied to a spacecraft configuration that uses solar-powered electric propulsion, including what might be referred to as a "fan-shaped" solar cell array. The fan-shaped panels represent a deliberate design attempt to avoid the very low natural frequencies (<0.1 Hz) and low damping ($<0.5\%$) associated with very long and narrow panels of either the fold-out or roll-out variety.

Using a detailed modal analysis of the panel structures *and* a large, high-gain antenna as well, a somewhat simplified version of the equations (described in SPS 37-54, Vol. III, pp. 65–69) was programmed for the IBM 7094 digital computer in DSL/90–Fortran IV and for an Electronic Associates 231-R analog computer. Due to the relative stiffness of the ion engine translator and gimbal actuators ($\omega_\eta > 100$ rad/s), their dynamic effects were deleted from the system model so as not to inflict these high frequencies unnecessarily on numerical integration routines. Momentum wheels were not included in the initial simulations. The resulting dynamical equations are

$$\ddot{\eta} + \Delta\ddot{\theta} = -2\zeta\omega\dot{\eta} - \omega^2\eta - \Phi^T M \Sigma_E F/\mathcal{M}$$

$$I\ddot{\theta} + \Delta^T\ddot{\eta} = T$$

or, in combined matrix form,

$$\begin{bmatrix} I & \bar{\Delta}^T \\ \bar{\Delta} & E \end{bmatrix} \begin{Bmatrix} \ddot{\theta} \\ \ddot{\bar{\eta}} \end{Bmatrix} = \begin{Bmatrix} T \\ -2\bar{\zeta}\bar{\omega}\dot{\bar{\eta}} - \bar{\omega}^2\bar{\eta} - \bar{\Phi}^T M \Sigma_E F/\mathcal{M} \end{Bmatrix} \tag{1}$$

where $\bar{\Delta} = -\bar{\Phi}^T M (\Sigma_E \tilde{R} + \tilde{r}\Sigma_E)$, $N \times 3$ (N = number of modal coordinates) and the overbars are adopted to indicate that the set of modal coordinates $\{\eta\}$ has been truncated to simplify the model and, more importantly, relieve the computational load as well.

The square matrix in Eq. (1) is constant and need only be inverted once to achieve an explicit solution for $\ddot{\theta}$ and $\ddot{\bar{\eta}}$

(of course, the right-hand side must be evaluated at each integration step). Values of I, $\bar{\Delta}$, $\bar{\zeta}$, $\bar{\omega}/2\pi$, and $\bar{\Phi}^T M \Sigma_E$ developed for the spacecraft configuration simulated here are given in Tables 1, 2, and 3. Total spacecraft weight was 2390 lb. Using the control system parameters developed by Mankovitz[7] with some gain changes, spacecraft motion was simulated under the conditions of initial tumbling rates amounting to 0.01 mrad/s about each axis (pitch, yaw, and roll) and, in effect, suddenly applied ion engine thrust (0.02 lb by each of 4 engines). The ion engine bank has also been rotated to a position 45 deg from the spacecraft's x- and z-axes (in the x–z plane) to test the coupling thus produced between translator and gimbal control loops.

Table 1. Modal frequency and damping

Mode	Frequency $\bar{\omega}/2\pi$, Hz	Damping factor $\bar{\zeta}$
1	0.60861	0.01
2	0.93037	0.01
3	1.0360	0.01
4	1.2616	0.01
5	1.6228	0.01
6	1.9434	0.01

Table 2. Undeformed vehicle inertia matrix (slug-ft²)

13360	39.2	992
39.2	9438	6.20
992	6.20	7270

Table 3. Rigid–flexible body coupling

Mode	$\bar{\Phi}^T M \Sigma_E$, slugs			$\bar{\Delta}$, slug-ft		
	x	y	z	x	y	z
1	3.4629	−0.0314	0.5928	−0.5654	−79.94	−0.0778
2	−0.001	0.1169	−0.0002	71.23	0.0039	11.65
3	−0.7024	0	3.962	−0.0059	0.0297	0.0397
4	0	−0.0033	0	−0.188	−0.0001	1.223
5	−3.519	0.0089	−0.5992	0.1539	21.60	−0.0131
6	−0.001	−0.651	−0.0004	−19.27	0.0001	81.16

Figures 21 through 25 indicate the simulated response of the system to engine startup and the subsequent rate reduction. The transient vibrations of the solar panel–antenna structure combination are seen to couple noticeably *only* into yaw rotations of the spacecraft bus.

[7] *1975 Jupiter Flyby Mission Using a Solar Electric Propulsion Spacecraft*, Mar. 1, 1968 (JPL internal document).

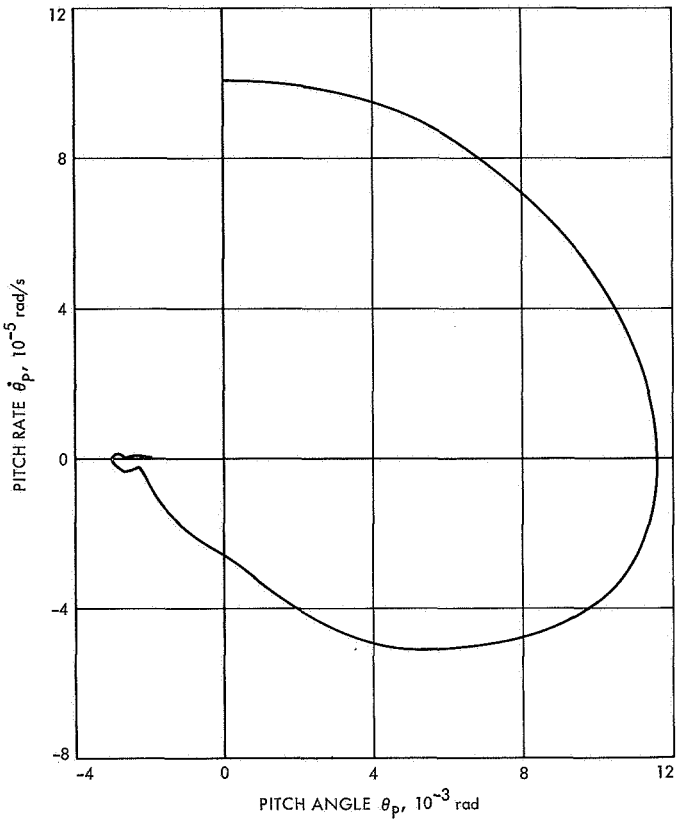


Fig. 21. Spacecraft pitch axis response

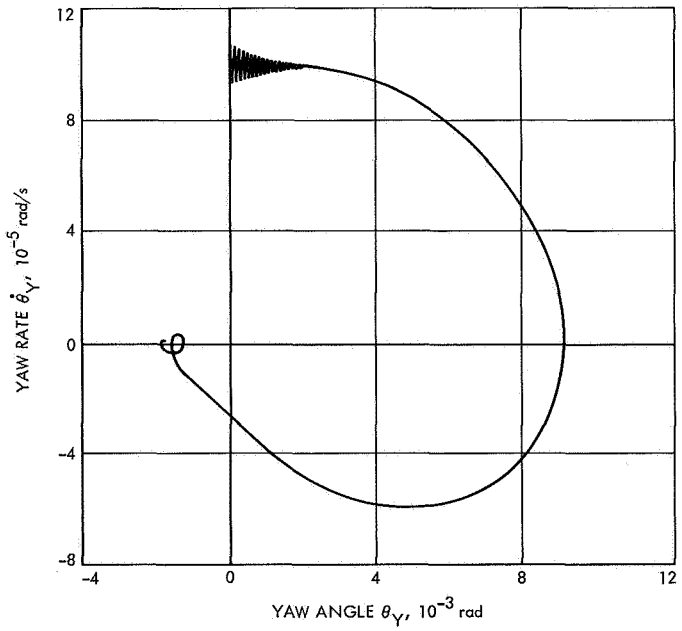


Fig. 22. Spacecraft yaw axis response

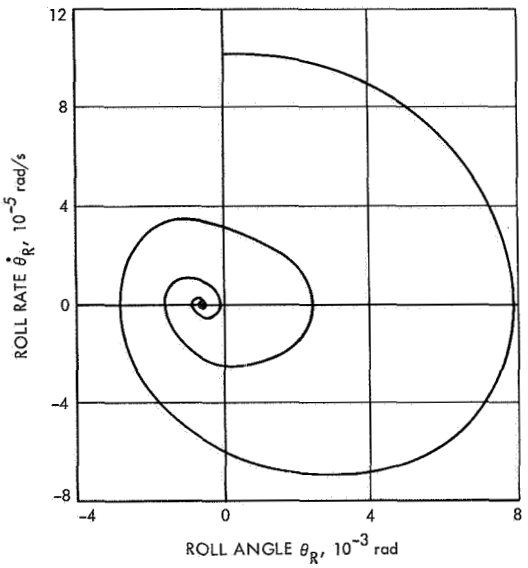


Fig. 23. Spacecraft roll axis response

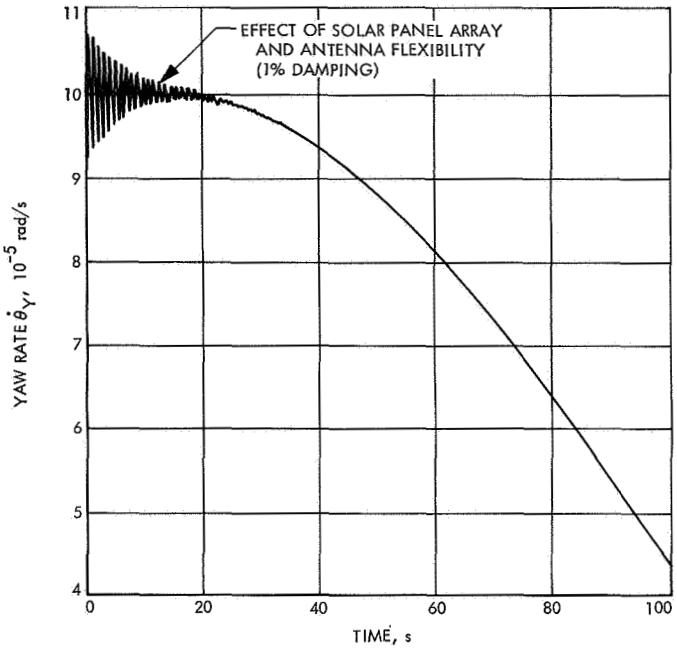


Fig. 24. Spacecraft yaw rate vs time

Further, it is apparent that the vibrations have no significant effect on the performance of the thrust vector–attitude-control system, so widely disparate are their characteristic response frequencies. Control system bandwidth extends to less than 0.01 Hz, while the natural structural vibrations occur at about 0.6 Hz or greater. Also, when the ion engine translator and gimbals are driven by the stepper motor, the step size and resulting reaction forces and torques are such that negligible excitation of the flexible structure occurs, compared with the engine startup transient. For this reason and for computational efficiency, the simulation of the engine control system was accomplished using a continuous actuator model rather than a discrete one.

Thus, it is clear that an even more flexible solar panel configuration could be accommodated without any problem from an attitude-control viewpoint. If the control system design were to widen its bandwidth, and if very flimsy solar arrays of either the roll-out or fold-out variety were to be used, another examination of control–structural dynamics interaction would be warranted. However, it appears that, at least for the solar electric configurations under study, spacecraft flexibilities will have their major impact on high-gain antenna pointing accuracies, etc., and have little to do with attitude-control performance or stability.

Reference

1. Likins, P. W., and Gale, A. H., "Analysis of Interactions Between Attitude Control Systems and Flexible Appendages," paper presented at the 19th International Astronautical Congress, Oct. 14–19, 1968 (*Proceedings* to appear).

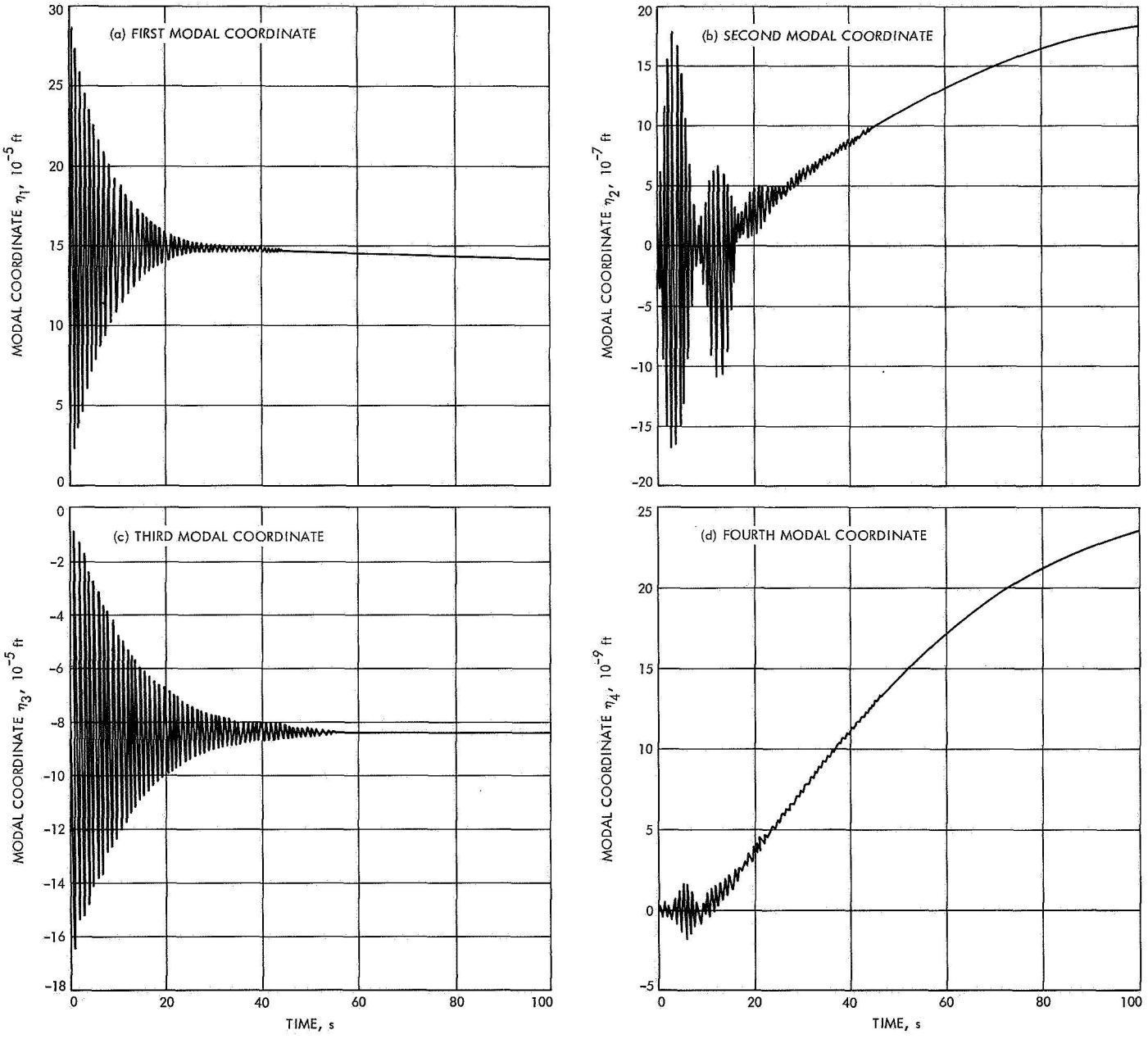


Fig. 25. Modal coordinates of spacecraft flexible appendage model

XI. Guidance and Control Research

GUIDANCE AND CONTROL DIVISION

A. Guard-Ring Thermionic Diode Experiments,

K. Shimada and P. L. Cassell

1. Introduction

During the past two years, we have been investigating means for improving the electrical output of thermionic energy converters by reducing electron-scattering losses in the cesium plasma. It was demonstrated that the plasma-maintenance voltage could be reduced at reduced cesium pressure with the aid of an external magnetic field (SPS 37-45, Vol. IV, pp. 42–45). To better understand this reduction, we are making a quantitative analysis of cesium ionization in a thermionic diode at a low cesium pressure.

Since ionization occurs at thermionic emitters by surface processes and in the interelectrode gap by electron-impact processes, and since these processes are controlled by electrode work functions, an accurate knowledge of these values is essential for quantitative analysis. However, the determination of emitter work functions is difficult in practical thermionic converters because a large current originating from the side walls adds to the current from the emitter whose work function is to be determined. Therefore, a diode with a guard ring but with dimensions similar to those of practical converters was constructed. With this diode we have determined: (1) emitter and collector work functions, (2) side-wall currents that closely represent those in practical converters, and (3) thermal characteristics of the diode.

2. Guard-Ring Diode[1]

The guard-ring diode (Fig. 1) has a planar electrode geometry with an interelectrode gap of 0.010 in. (0.254 mm). The collector, which is surrounded by a guard-ring electrode, has an area of 1.53 cm^2 and therefore the defined emitter area equals 1.53 cm^2 when the guard-ring is used properly. Niobium is used for both the collector and the guard ring, and tantalum for both the emitter and the emitter heat choke.

Two salient features in the construction of the diode are the guard ring and the calorimeter. The guard ring is a thin-walled cylinder, fitted around the collector barrel, electrically insulated from it by a thin (0.010 in. = 0.254 mm) alumina layer. Because of this construction, the

[1]Fabricated for JPL by the Thermo Electron Co., Waltham, Mass.

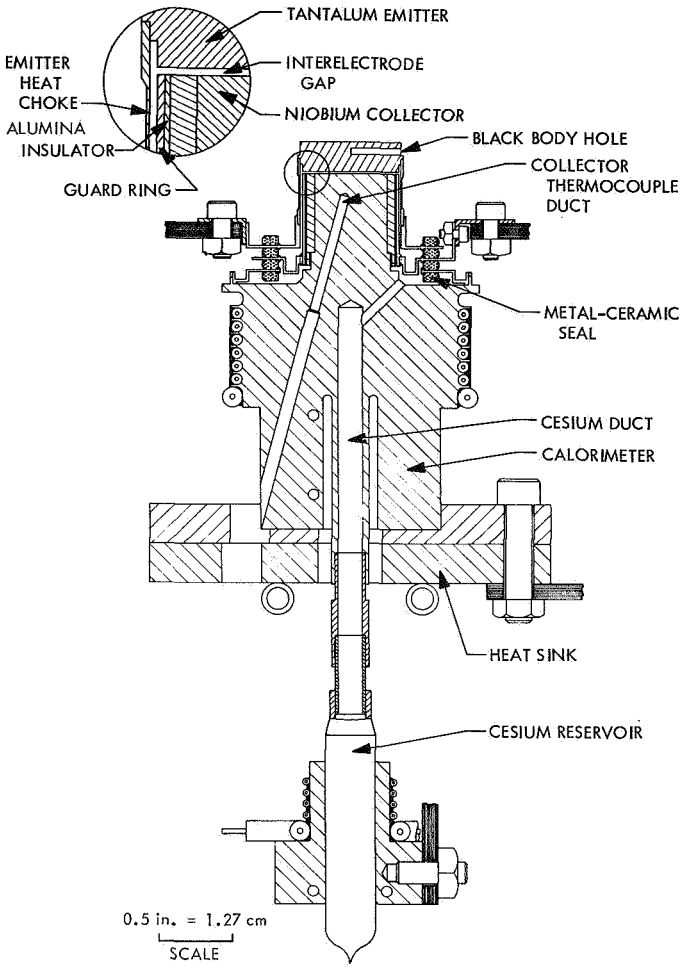


Fig. 1. Guard-ring research converter

temperature of the guard ring remains nearly equal to that of the collector without the separate heating required for conventional guard-ring diodes. The surface of the guard ring, which is parallel to that of the emitter, is flush with the collector surface. The cylindrical sides of the guard ring and emitter heat choke are parallel and separated by a narrow gap (0.015 in. = 0.381 mm), whereas, in contrast, the emitter heat choke faces the cylindrical sides of the collector in practical energy converters. Consequently, the guard ring of this research converter can be used to estimate the side-wall currents in practical converters.

Another unique feature of this diode is the calorimeter section that provides for heat-flux measurements. This section is located between the collector heater and the heat sink, and has two thermocouples for measuring temperature gradients. Since the dimensions are accurately known, the heat flux can be determined from the measured temperature gradient and the known thermal properties of the rhenium calorimeter.

The diode is also equipped with a copper block attached to the cesium reservoir to reduce temperature errors at the liquid–vapor interface, whose level depends on the cesium temperature. The emitter has a black-body hole (depth/diameter $\simeq 8$), which is drilled in parallel to and separated by 0.13 in. (3.3 mm) from the emitter surface. Another hole is drilled into the collector barrel to accommodate a deep immersion thermocouple for collector temperature measurements. The top of the thermocouple lies approximately 0.2 in. (5 mm) from the collector surface. The collector surface temperature was accurate within 5°K for all values of heat flux encountered.

3. Measurements

The diode was operated in an ion-rich, unignited mode to determine its electrode work functions. To avoid electron scattering, the cesium pressure was maintained in a range in which electron-neutral atom collisions are minimal.

Volt–ampere curves of the diode, plotted with an x–y recorder, display directly the diode current as a function of the voltage measured at the potential terminals of the diode, which are located on its emitter and collector. The display may be in either linear or semilog form.

For each measurement, the temperatures of the cesium reservoir and the collector were held within 0.5 deg of preset values by two temperature controllers. These temperatures were measured by uninterrupted chromel–alumel thermocouples referenced to 0°C with the electronic reference junctions. The emitter temperature was measured with an optical pyrometer, and was kept constant within 3 deg by controlling the bombarding power output of an electron-gun supply.

Emitter work functions were determined from the saturation-electron-current density obtained from volt–ampere curves. During these measurements, the electric potential of the guard ring was electronically equalized to that of the collector for its entire sweep from -5 to $+5$ V.

The equalization was accomplished by offsetting the guard-ring potential until the leakage current, through a circuit including guard-collector leakage resistance ($\simeq 1\ \Omega$) and the current shunt, became zero. The required offset was approximately 1 mV, and its polarity was such that the parasitic thermal emf in the loop was cancelled out.

When equalization was achieved, the vacuum levels (Fermi level + work function) were also equalized, since the electrodes were both made of niobium and held at nearly equal temperatures.

Emitter work functions determined from the saturation currents of the volt–ampere curves for a tantalum emitter are plotted against temperature ratio T_E/T_{Cs} in Fig. 2. The cesiated true work function is seen to equal that of a material with an uncesiated work function ϕ_0 of 4.7 V. This value was higher than expected for tantalum, and might have resulted from oxygen impurities. The apparent emitter work function was also determined from the total current from the emitter and its supporting structure to the collector and its guard ring. The saturation current of the diode was determined from the short-circuit current occurring at zero volts across the diode, at which any leakage current was also zero. Since the diode current actually saturated at voltages less than zero, and since the saturation current increased negligibly as the diode voltage increased, the current determined at zero volts practically equals the saturation current. The apparent emitter work function thus determined is also shown in Fig. 2. Note that it is less than the true emitter work function by as much as 0.2 V. The observed difference in work functions is attributed to large side-wall currents which were added to the true emitter current during the measurement of apparent work functions. Total current thus measured was equivalent to that from an emitter whose work function was lower than the true work function. From the true work functions, heat-choke temperatures, and dimensions, side-wall currents were calculated by using a theory developed earlier (SPS 37-52, Vol. III, pp. 65–68). Subsequently, apparent emitter work functions were deduced from the true emitter current plus the calculated side-wall currents. The results shown by closed points in Fig. 2 along the measured apparent work function indicate the applicability of the theory for calculating side-wall currents.

Collector work functions were obtained from: (1) emitter current measurements in the retarding voltage range of the volt–ampere curve, and (2) current measurements in the back-emission mode in which electrons were emitted mainly from the collector. Typical retarding plots are shown in Fig. 3. The collector work functions were determined from these curves by using the relationship

$$\ln J_C = \ln (A T_E^2) - e(\phi_C + V)/kT_E$$

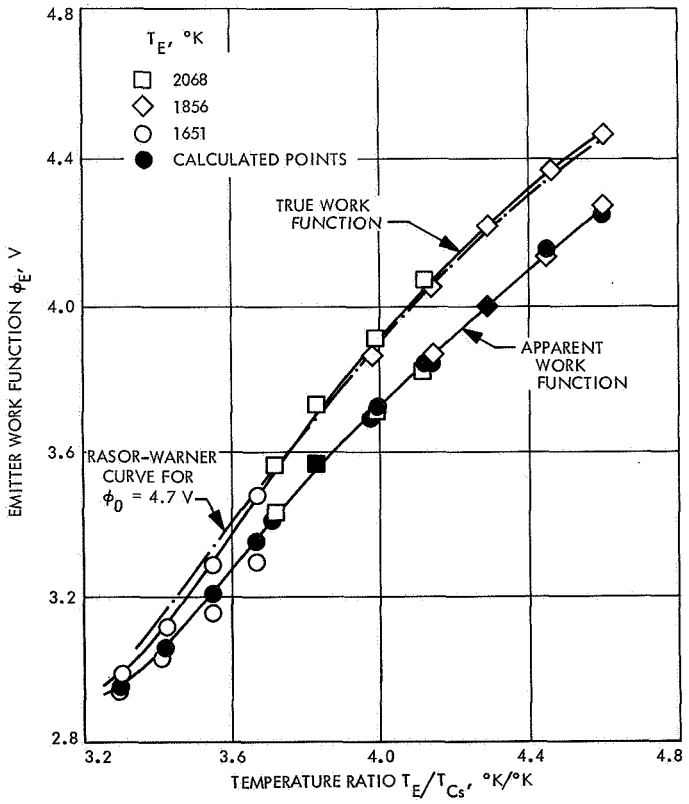


Fig. 2. Emitter work function vs temperature ratio

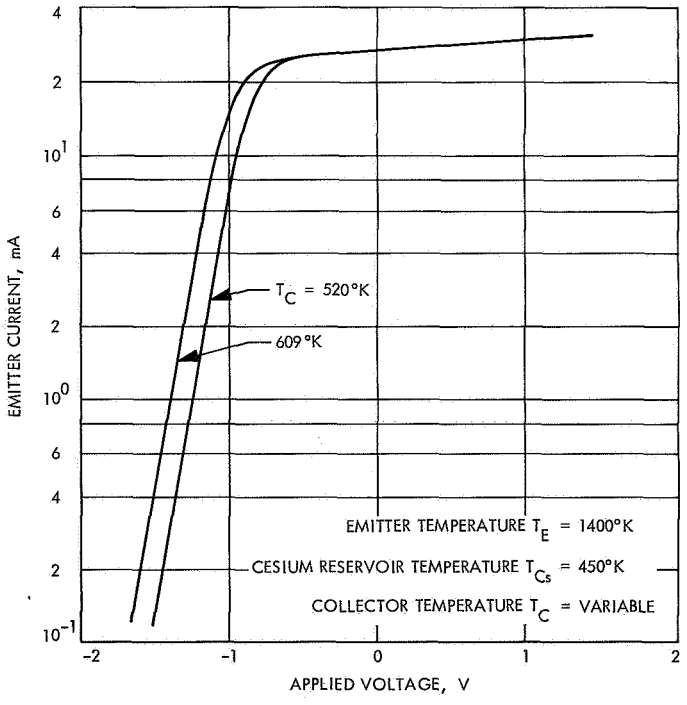


Fig. 3. Volt–ampere curve in retarding voltage range

where

J_C = collected emitter current density, A/cm²

A = 120 A/°K-cm²

T_E = emitter temperature, °K

$e = 1.6 \times 10^{-19}$ C

$k = 1.38 \times 10^{-23}$ J/°K

ϕ_C = collector work function, V

V = output voltage (negative applied voltage) corresponding to current J_C, V

The calculated work functions are shown in Fig. 4 for T_C/T_{Cs} less than 1.7. Scatter in the results as large as ±50 mV reflects the difficulties in such measurements. Meaningful results can be obtained only when the electron space charge near the collector is neutralized, and when the net ion current from the emitter and the back-emission current from the collector are both negligible in comparison with the electron emission from the collector. Consequently, temperature ratios T_C/T_{Cs} for these measurements were restricted to the range 1.1–1.7. For values of T_C/T_{Cs} larger than 1.7, the back-emission method was used. Again, difficulties were experienced in obtaining consistent results for collector work functions. The results are also shown in Fig. 4, together with a typical work function of a molybdenum collector. Present results for a niobium collector show that its work function is higher than that of a molybdenum collector by approximately 0.1 V. (No published data are available for niobium collector work functions.) Difficulties arising from the space-charge limitation that occurred for large temperature ratios could have been reduced if the interelectrode spacing were adjustable to a smaller value than our present 0.010 in. (0.254 mm).

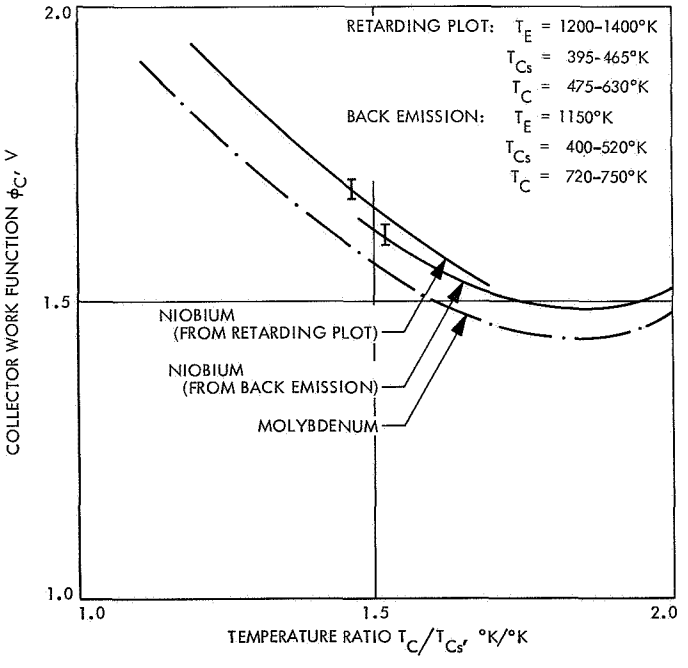


Fig. 4. Collector work function vs temperature ratio

The temperatures of the collector surface were estimated from those measured by an immersion thermocouple located 0.2 in. (5 mm) from the surface. Differences in these temperatures were calculated from the geometry and from heat fluxes which, in turn, were determined from the calorimeter. The analysis showed that the heat flow by conduction through metal–ceramic seals was much larger than that by radiation to the collector from the emitter at those temperatures suitable for measurements of collector work functions. The largest temperature difference was estimated to be 5°K, and therefore the errors in work-function measurements due to temperature errors would have been 20 mV.

4. Conclusion

In a guard-ring thermionic diode operating in an unignited mode, side-wall currents as large as the emitter current were measured. The diode had dimensions typical of thermionic energy converters with plane-parallel electrode geometry. It was also found that the side-wall currents could be calculated quite accurately by using a theory developed earlier (SPS 37-52, Vol. III).

The niobium collector work functions were measured by a retarding-plot method for temperature ratios T_C/T_{Cs} between 1.1 and 1.7. For larger ratios, back-emission measurements were employed. Formation of an electron-space-charge sheath that modified the measured collector work function often resulted in unreasonably large values. The minimum work function for this niobium collector was 1.5 ± 0.05 V occurring at $T_C/T_{Cs} \simeq 1.8$. Although there are no published data for niobium collector work functions, our value appears reasonable in comparison with those of similar refractory metals including molybdenum in cesium environment.

Work functions measured in the guard-ring diode will be subsequently incorporated in the quantitative analysis of pre-ignition characteristics in which cesium ionizations at electrode surfaces and in the interelectrode space occur simultaneously.

B. Thermal Noise of Injected Carriers, M-A. Nicolet[2]

1. Introduction

The experimental facts on the high-frequency noise of single and double injection devices indicate that the mean square of the voltage fluctuations is different in the two cases. For single injection, it seems that (Refs. 1 and 2)

$$\langle v^2 \rangle \approx 2 \cdot 4kTr\Delta f \tag{1}$$

but

$$\langle v^2 \rangle = 1 \cdot 4kTr\Delta f \tag{2}$$

for double injection (Ref. 3), where r is the differential resistance (real part of the impedance) of the device at the frequency under consideration. These expressions are of the form

$$\langle v^2 \rangle = \alpha \cdot 4kTr\Delta f \tag{3}$$

with $\alpha \approx 2$ for single and $\alpha = 1$ for double injection. This article shows that:

(1) The model of thermal random noise proposed by Webb and Wright (Ref. 4) and modified by van der Ziel (Refs. 5 and 6) predicts such a difference.

(2) α acquires varying values, depending on the conditions prevailing in the device and on the frequency ranges these conditions establish by virtue of their natural time constants.

The predictions of this model have been derived to determine which experiments would be best suited to test the model and its underlying hypothesis.

2. Model

According to this "thermal hypothesis" the random motion of charge carriers injected into the bulk of a solid generates a mean square voltage fluctuation across a volume element $A \cdot \Delta x$ given by

$$\Delta \langle v^2 \rangle = 4kT\Delta r\Delta f \tag{4}$$

and the total thermal noise across the bulk is obtained by the summation

$$\langle v^2 \rangle = 4kT\,(\Sigma\Delta r)\,\Delta f \tag{5}$$

where

$$\Delta r = \Delta x/A\rho\mu \tag{6}$$

and ρ is the charge carrier concentration and μ its mobility. We now assume that:

(1) The current is due to drift only.

(2) The contribution of the contacts to the noise is negligible ($\rho \to \infty$).

It then follows that $\langle v^2 \rangle$ across the device must be given by

$$\langle v^2 \rangle = 4kT\,(V/I)\,\Delta f \tag{7}$$

where V and I are the dc values of the voltage and current at the operating point. Comparing Eq. (3) with Eq. (7), one finds that

$$\alpha = \frac{V/I}{r} \tag{8}$$

The connection between V/I and r depends on the mechanism of operation of the device. Several limiting cases will be distinguished according to: (1) the number of carrier types contributing to the current, and (2) the position of some characteristic time τ of the device with respect to the time period $t = 2\pi/\omega$ of the frequency ω at which the noise is measured.

3. Single Injection

In a single-carrier space-charge-limited current (SCSCLC) device, the following natural time intervals exist:

τ_0	transit time of carrier from emitter to collector
$\langle \tau \rangle$	mean "free time" of the carrier (mean time between collisions)
τ_{rel}	dielectric relaxation time, formally definable as $\epsilon/\mu\,\langle \rho \rangle$, where $\langle \rho \rangle$ is some suitable average of ρ

Typical values of $\langle \tau \rangle$ for high-purity silicon at room temperature are 10^{-12} to 10^{-14} s. Typical values of $\langle \tau_0 \rangle$ for a planar configuration are given by the formula $\tau_0 = (4/3)\,L^2/V\mu$, where L is the length of the device. A correct order of magnitude of τ_{rel} is found by setting $\langle \rho \rangle$ equal to the value of ρ at the collector, $(4/8)\,\epsilon V/L^2$. This yields $\tau_{rel} \approx (8/4)L^2/V\mu \approx \tau_0$. There are thus only two characteristic times in single injection devices, the mean free time $\langle \tau \rangle$ and the transit time τ_0.

[2]At the California Institute of Technology, performing work supported by JPL and NASA Electronics Research Center, Cambridge, Mass.

Since the carriers are assumed to be "thermalized" in their motion through the space-charge region, the condition $\langle\tau\rangle < t$ must obviously be imposed. We therefore consider only the two remaining ranges of operation, the high-frequency range $\langle\tau\rangle < t < \tau_0$ and the "low"-frequency range $\langle\tau\rangle < \tau_0 < t$.

In the low-frequency range $\tau_0 < t$, the differential resistance r is equal to the quasi-static value $\partial V/\partial I$. If the I–V characteristic has the form $I \propto V^n$, then $r = (1/n)\,V/I$ and by Eq. (8)

$$\alpha = n, \qquad (I \propto V^n, \tau_0 < t) \tag{9}$$

This result was already stated in the corresponding form $\langle i^2\rangle = n \cdot 4kTg\Delta f$ by Yamamoto, et al. (Ref. 7). Experimental evidence in support of this formula has been offered for devices of GaAs, CdS, Ge and Si and $n = 2$. The results obtained on CdS by Liu, et al. (Ref. 2) are the most conclusive ones so far and agree well with Eq. (9).

The agreement for $n = 2$ is to be expected because this case corresponds to space-charge-limited current of thermal charge carriers (carriers with constant mobility) through structures of almost arbitrary geometrical configuration (Ref. 8). Departures from $n = 2$ usually indicate that other effects are present, such as heating of lattice and/or carriers, trapping, surface leakage, etc. The generalization of Eq. (9) to values other than $n = 2$ will therefore require both theoretical and experimental justification.

In the high-frequency range $\langle\tau\rangle < t < \tau_0$, the differential resistance r has the value $(3/2)\,\partial V/\partial I$ for an SCSCLC device of planar geometry and $n = 2$ (Ref. 9). Thus,

$$\alpha = 4/3, \qquad (I \propto V^2, \langle\tau\rangle < t < \tau_0) \tag{10}$$

How r relates to $\partial V/\partial I$ for structures other than planar is not known.

Noise measurements in this range of frequencies have never been performed on SCSCLC. Equation (10) predicts that the noise performance of such devices improves, but only slightly, at high frequencies.

4. Double Injection

Double injection generally refers to those cases in which both electrons and holes play a role in the conduction mechanism. This does not necessarily imply that electrons and holes both carry current in non-negligible proportions. Even when it is so, the distinct cases of double injection possible are still numerous. This discussion is limited to simple cases to demonstrate the principle through examples.

The natural time intervals of a double injection (two-carrier) space-charge-limited current (TCSCLC) device are:

τ_n, τ_p	lifetime of electrons and holes
τ_{0n}, τ_{0p}	transit time of electrons and holes
$\langle\tau_n\rangle, \langle\tau_p\rangle$	mean "free time" of electrons and holes (mean time between collisions)
τ_{rel}	dielectric relaxation time, formally definable as $\epsilon/\langle\mu\rho\rangle$, where $\langle\mu\rho\rangle$ is some suitable average of the conductivity

The lifetimes τ_n, τ_p are always the longest and $\langle\tau_n\rangle, \langle\tau_p\rangle$ are normally the smallest.

a. Semiconductor regime. When the concentrations δn and δp of injected electrons and holes are large compared to the thermal equilibrium values n_0 and p_0 of mobile charge carriers and the fixed charge concentration $p_0 - n_0$ of the doping (but $\delta p - \delta n$ is still small compared to $p_0 - n_0$), the I–V characteristic of a long double injection diode of planar configuration, length L and area A, is (Refs. 10 and 11)

$$I = A\,(9/8)\,q\mu_n\mu_p\,(p_0 - n_0)\,\tau V^2/L^3 \tag{11}$$

where τ is the (common) lifetime of the injected electrons and holes and μ_n and μ_p are their mobilities. In this "semiconductor" regime, the transit time of the carriers is given by $\tau_{0n} \approx L^2/\mu_n V$ and $\tau_{0p} \approx L^2/\mu_p V$ and $\tau_{0n}, \tau_{0p} << \tau$, to assure that double injection conditions prevail. The condition that the net space charge $q\,(\delta p - \delta n)$ is much smaller than the injected charge leads to the requirement $\tau_{\mathrm{rel}} << \tau_{0n} \cdot \tau_{0p}/(\tau_{0n} + \tau_{0p}) < \tau_{0n}, \tau_{0p}$. In this case, therefore, the relevant time constants are, in increasing order, $\langle\tau_n\rangle$ & $\langle\tau_p\rangle$, τ_{rel}, τ_{0n} & τ_{0p}, and τ. From the thermal hypothesis $\langle\tau_n\rangle, \langle\tau_p\rangle < t$, we thus distinguish four distinct ranges of operation. To establish the relative significance of these ranges, we apply a differential voltage step ΔV across the device and follow the current response.

(1) $\langle\tau_n\rangle$ & $\langle\tau_p\rangle < t < \tau_{\mathrm{rel}}$. The additional field generated by ΔV is constant throughout the length of the device and has the value $\Delta E = \Delta V/L$. The additional current ΔI

is caused by the drift of the steady-state carrier concentration before the application of the pulse. This concentration is

$$\rho = [(3/4)\, q\, (p_0 - n_0)\, \tau\mu_n\mu_p V/(\mu_n + \mu_p)\, L^2]\, (L/x)^{1/2} \tag{12}$$

and yields

$$r = (3/4)\, V/I \tag{13}$$

hence,

$$\alpha = 4/3 \tag{14}$$

as in the corresponding case of SCSCLC.

(2) $\tau_{rel} < t < \tau_{0n}$ & τ_{0p} and τ_{0n} & $\tau_{0p} < t < \tau$. The drift current generated by the constant electric field $\Delta V/L$ and the non-constant carrier distribution of Eq. (12) is not solenoidal. This leads to a buildup of net local charge and a relaxation of the field, but still without significant change in the carrier concentration. The new field distribution $\Delta E\,(x)$ now generates a divergence-free current ΔI everywhere. Since the charge carrier distribution is unmodified as yet, $\Delta E\,(x)$ is proportional to the distribution of the steady-state field $E\,(x)$ prior to the application of the pulse. Hence, $(\Delta E/E) = (\Delta V/V) = (\Delta I/I)$, and therefore

$$r = \frac{\Delta V}{\Delta I} = \frac{V}{I} \tag{15}$$

so that

$$\alpha = 1 \tag{16}$$

Although free of divergence, the new drift current still is not in equilibrium with the carrier generation-recombination (g-r) process in a volume element. This equilibrium is achieved (over many transit times) with a time constant of τ. Only the lifetime τ, therefore, marks a change in the behavior of the system.

(3) $\tau < t$. In this range of quasi-steady-state

$$r = \partial V/\partial I = (1/2)\, (V/I) \tag{17}$$

and

$$\alpha = 2 \tag{18}$$

An experimental investigation of a Si double injection device operated in the semiconductor regime has recently been made by Bilger et al. (Ref. 3). They observed the noise in the range $\tau_{rel} < t < \tau$ and found $\alpha = 1.04 \pm 5\%$, in agreement with Eq. (16). At frequencies below $f = 1/\tau$, an additional component of noise, interpreted as g-r noise, strongly dominates over thermal noise. No results are available at frequencies beyond $f = 1/\tau_{rel}$.

b. A general case. Some generalizations of the previous results are obvious. If a TCSCLC device has a dc characteristic of the form $I \propto V^n$ (in contrast to SCSCLC, this is possible even with thermal carriers), the thermal noise will be given by

$$\alpha = n \tag{19}$$

for frequencies in the range of quasi-steady-state (up to the lowest recombination frequency $1/\tau$). However, a component of g-r noise is also simultaneously expected there and will, in general, dominate.

At frequencies beyond the fastest recombination frequency $1/\tau$, on the other hand, the electrons and holes are essentially decoupled; the device represents a parallel combination of the two conductances presented by the electron and hole distribution separately. Each generates thermal noise; hence,

$$\alpha = 1 \tag{20}$$

up to the dielectric relaxation time. The device also behaves resistively for differential signals in that range.

To evaluate α at frequencies beyond $1/\tau_{rel}$, the steady-state distribution of the carrier concentrations, or that of the electric field E, must be known. The value of α in this range is then readily found to be

$$\alpha = L^{-2}\left(\int_0^L E\, dx\right)\left(\int_0^L E^{-1}\, dx\right) = (V/L)\, (\mu_n + \mu_p)\, \langle\rho\rangle/j \tag{21}$$

indicating that at those frequencies noise depends only on the average value of $\langle\rho\rangle$ of the carrier concentration. The value of $\alpha = 4/3$ in Eq. (14) follows from this general formula for the special case of $E \propto x^{1/2}$.

c. Insulator regime. In this mode of double injection, the net space charge $q(\delta p - \delta n)$ is no longer small compared to $q(p_0 - n_0)$, as in the semiconductor regime; rather $p_0 - n_0 < \delta p - \delta n$. The net space charge, however, is still much smaller than the injected charge, so that the relative positions of the various time constants remain unaffected. The *I–V* characteristic, on the other hand, changes to (Refs. 10 and 11)

$$I = A\, (125/18)\, \epsilon\mu_n\mu_p\tau V^3/L^5 \tag{22}$$

for a planar configuration of the device, and the carrier distribution becomes symmetrical with respect to the two contacts. A detailed calculation establishes that for these conditions

$$\langle \rho \rangle = (6/5)\, jL/(\mu_n + \mu_p)\, V \tag{23}$$

From the preceding section, one derives the following values of α for the various ranges:

$$\left.\begin{array}{lll} \alpha = 3 & \text{for } \tau < t & \text{(quasi-static range)} \\ \alpha = 1 & \text{for } \tau_{rel} < t < \tau & \text{("linear" intermediate range)} \\ \alpha = 6/5 & \text{for } \langle \tau_n \rangle, \langle \tau_p \rangle < t < \tau_{rel} & \text{(high-frequency range)} \end{array}\right\} \tag{24}$$

5. Concluding Remarks

The present method of identifying time intervals, in which the physical processes at work can naturally be divided to determine the thermal noise through the differential resistance r and then relate it to V/I, is generally applicable. If trapping is present, for instance, and the time constants for the process are all large compared to the dielectric relaxation time, there must always be a range of thermal noise with $\alpha = 1$ in double injection. Conversely, a value of $\alpha \neq 1$ over a given frequency range indicates the presence in the corresponding time interval of other physical processes superposed on the thermal equilibrium of the charge carriers, or even violating it.

The results derived here from the thermal noise hypothesis as proposed by van der Ziel seem to agree with the experimental facts, where observations are available. These observations are presently limited to two distinct cases: SCSCLC with $\alpha = 2$ and TCSCLC with $\alpha = 1$. The latter case is not specific as a test of the model, because Webb and Wright's original proposal of $\langle v^2 \rangle = 4kTr\Delta f$ predicts the same result. A real test on the diverging theories is provided only when $\alpha \neq 1$, i.e., for noise in SCSCLC and in TCSCLC at high frequencies (the low-frequency range is dominated by g-r noise). Table 1 summarizes this comparison.

The easiest test of the thermal hypothesis is thus offered by the noise of SCSCLC in a device with thermal carriers ($I \propto V^2$). The results obtained on CdS which are reported in Refs. 1 and 2 agree quite well with $\alpha = 2$. Unfortunately, no noise spectra are given there. A review of all cases of SCSCLC noise found in the literature also reveals that the information provided on the devices is never sufficient to clearly establish their mode of operation. *I-V* characteristics alone are known to be misleading at times. One therefore concludes that the experiment best suited

Table 1. Value of α predicted from van der Ziel's model[a]

SCSCLC		Low frequencies (quasi-static) $\tau_0 < t$	High frequencies $\langle \tau \rangle < t < \tau_0$	Very high frequencies $t < \langle \tau \rangle$
$I \sim V^2$ (constant mobility)	[b]	2	4/3	?
$I \sim V^n$ (a general case)	[b]	n (?) (needs justification)	?	?

TCSCLC		Low frequencies[c] (quasi-static range) $\tau < t$	Intermediate frequencies ("linear" range) $\tau_{rel} < t < \tau$	High frequencies $\langle \tau_n \rangle, \langle \tau_p \rangle < t < \tau_{rel}$	Very high frequencies $t < \langle \tau_n \rangle, \langle \tau_p \rangle$
$I \sim V^2$ (semiconductor regime)	[b]	2	1	4/3	?
$I \sim V^3$ (insulator regime)	[b]	3	1	6/5	?
$I \sim V^n$ (a general case)	[b]	n	1	Depends on $\langle \rho \rangle$ (Eq. 21)	?

[a] Predicted from van der Ziel's model (Refs. 5 and 6) when the thermal noise of injected carriers is written as $\langle v^2 \rangle = \alpha \cdot 4kTr\Delta f$. Webb and Wright (Ref. 4) propose $\alpha \equiv 1$ throughout.

[b] Expected to be dominated by excess noise at the lowest end of the frequency spectrum.

[c] Expected to be dominated by g-r noise.

to test the thermal hypothesis is to observe white noise levels, clearly free of excess noise, on devices of SCSCLC with thermal carriers ($I \propto V^2$), and independently identified as such by conclusive dc, ac and temperature measurements as well as by device design and fabrication data.

Finally, it is pointed out that the assumption of negligible diffusion is rather restrictive and excludes, for example, low injection junction devices.

References

1. Liu, S. T., *Solid-State Electron.*, Vol. 10, pp. 253–254, 1967.
2. Liu, S. T., et al., *Physica*, Vol. 38, pp. 279–284, 1968.
3. Bilger, H. R., et al., *J. Appl. Phys.*, Vol. 39, pp. 5913–5918, 1968.
4. Webb, P. W., and Wright, G. T., *J. Brit. IRE*, Vol. 23, pp. 111–112, 1962.
5. van der Ziel, A., *Solid-State Electron.*, Vol. 9, pp. 899–900, 1139–1140, 1966.
6. van der Ziel, A., *Solid-State Electron.*, Vol. 11, pp. 508–509, 1968.
7. Yamamoto, S., et al., *Appl. Phys. Letters*, Vol. 11, pp. 140–141, 1967.
8. Lee, D. H., and Nicolet, M-A., *Solid-State Electron.*, Vol. 8, pp. 182–184, 1965.
9. Baron, R., et al., *J. Appl. Phys.*, Vol. 37, pp. 4156–4158, 1966.
10. Lampert, M. A., and Rose, A., *Phys. Rev.*, Vol. 121, pp. 26–37, 1961.
11. Baron, R., *J. Appl. Phys.*, Vol. 39, pp. 1435–1446, 1968.

XII. Electronics Parts Engineering

ENGINEERING MECHANICS DIVISION

A. Power Pulse Method of Screening Resistors,

K. C. Evans

A machine has been purchased that will screen resistors by the power pulse method. This method applies a short duration power pulse to the part and the resistance variation is monitored during the pulse period. A maximum of 200 W is provided for resistors up to 20,000 Ω and a maximum of 2000 V for resistors between 20,000 Ω and 2.0 MΩ.

The machine is mounted in a two-bay console (Fig. 1) and consists of power supplies, power bridge, and control circuitry. Shown are the 0–70-V, 20-A dc power supply, control circuitry, null detector, power bridge, output recorder, null switch, test switch, test part holder clamp, two 0–220-V, 3-A power supplies connected in series for 440 V and the 0–2000-V, 0.5-A power supply. A foot switch controls the test part holder clamp.

The machine is used by first determining the amount of voltage required to supply the necessary power. The output recorder is calibrated to indicate a known change in resistance. After the proper power level is set up, the machine automatically plots the change in resistance during the power pulse. The shape of the output curve is used to determine the quality of the part under test. Each type of resistor material will produce output curves with different shapes.

Typical output curves for composition and metal film resistors are shown in Fig. 2. The composition resistor graph shows that, during the time power is applied, there is an initial decrease followed by an increase in resistance. Any composition resistor that deviates from the general shape could be cause for rejection. A metal film resistor will show an increase in resistance during the time the pulse is applied. All parts will show an initial offset on the graph, due to transients in the system, and can be ignored. The amount of resistance change or the variation from a normal trace that can be allowed without being considered a defective resistor varies depending on the type of resistance element and the degree of discrimination which is required.

The shape of the curve of a defective part will depend on the cause of the defect. Figure 3 illustrates a change in resistance versus time curve of a wirewound resistor

that has shorted turns. The shorted turns condition can be caused by adjacent turns touching each other or by a foreign particle touching several turns. The differential expansion in the resistor during power pulse testing causes intermittent contact between the turns of wire. Hence, the jagged and erratic traces.

Tests at JPL will further explore the meaning of the traces and the causes. Infrared techniques will be used to determine hot-spot location and the temperature of the part during pulse testing. Pulse level and duration will be explored to obtain optimum effectiveness of the power pulse method of screening resistors.

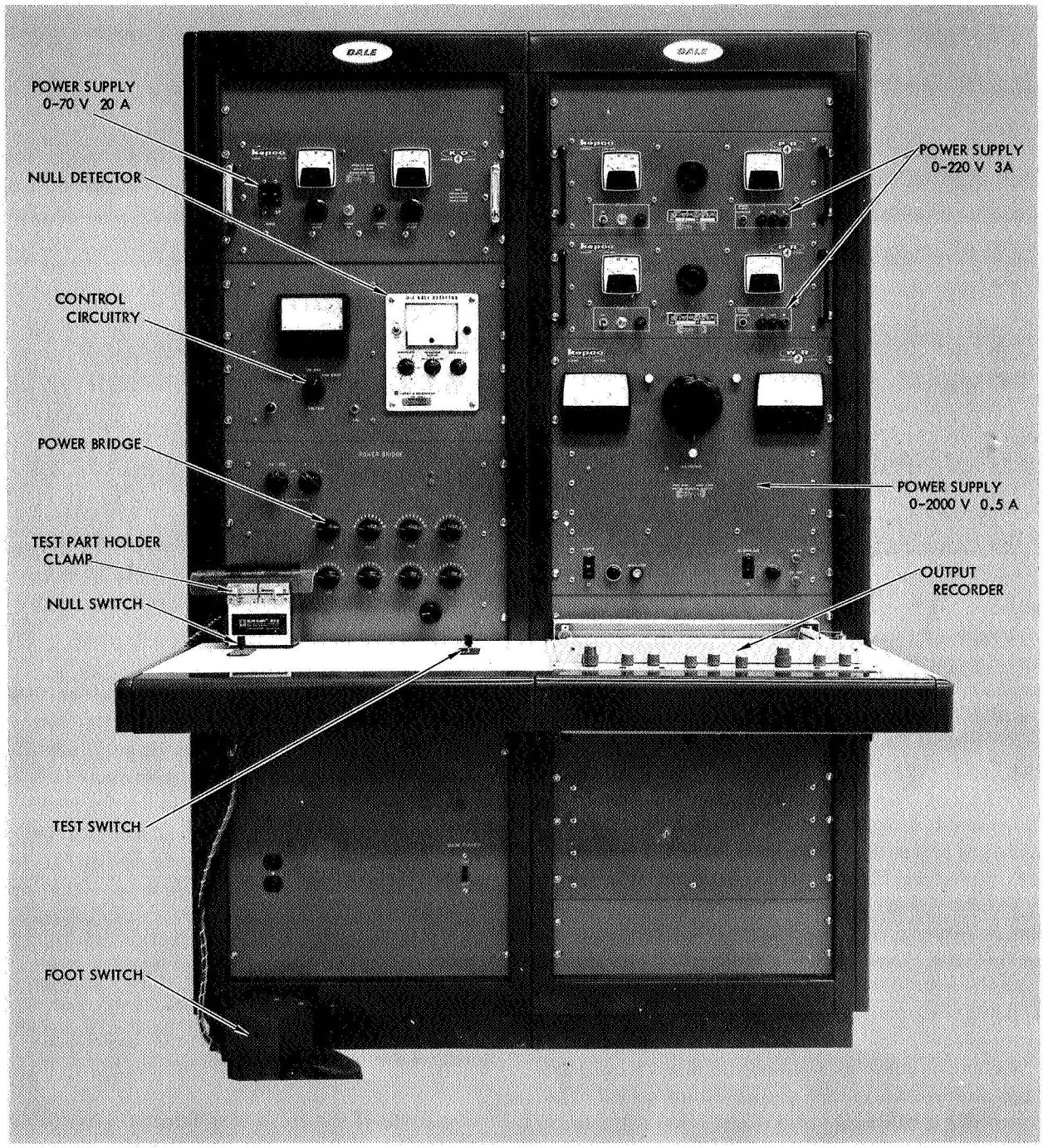


Fig. 1. Resistor power pulse tester

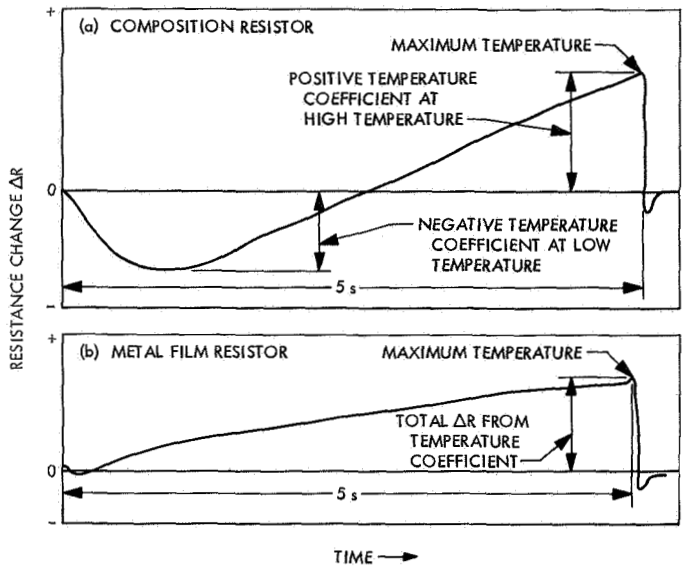


Fig. 2. Typical resistance traces

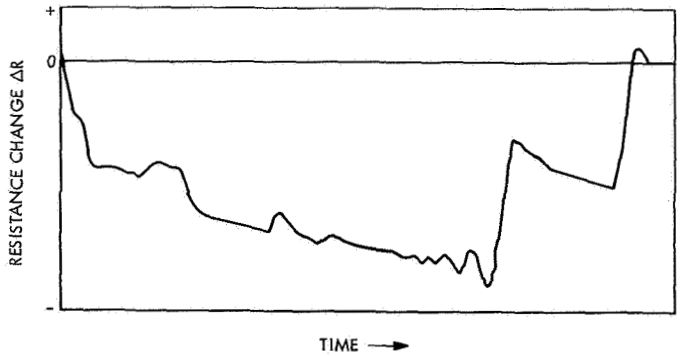


Fig. 3. Resistance change of wirewound resistor with shorted turns

XIII. Space Simulators and Facility Engineering

ENVIRONMENTAL SCIENCES DIVISION

A. Cryogenic Oscillators for Quartz Crystal Microbalances, *E. Laue and C. H. Goldsmith*

The use of vibrating quartz crystals to measure the mass of vacuum-deposited thin films has been developed into commercially-available equipment, and the mass production of multilayer films using quartz crystals as thickness monitors is a rapidly growing technical field. References 1 and 2 are concerned with the basic considerations for these applications. The major advantage of the quartz crystal microbalance technique lies in the fact that the relatively simple and portable equipment required has an inherent accuracy and resolution to 1 part in 10^7. Additional applications of the quartz crystal microbalances (i.e., chemisorbtion, biological interaction, corrosion measurement, and contamination detectors) will become available when controlled coating techniques are developed. Currently, cryogenic quartz crystal oscillators are being developed at JPL to monitor cryopumping, sublimation pumping, and contamination in the Space Molecular Sink Simulator Facility (commonly referred to as the Molsink facility). This article will concern the application problems associated with the Molsink environment.

Some details concerning the Molsink facility can be found in Ref. 3. The environmental parameters that affect the microbalance application involve a temperature range of 14 to 280°K, pressure from 10^{-12} to 760 torr, the requirement for minimum outgassing of materials in the test chamber, and the double-walled vacuum system, involving two vacuum feedthroughs, requiring a minimum separation of 6 m between the transducer and readout instrumentation. The primary difficulty concerns the operation of junction semiconductors in the oscillator circuit. At temperatures below approximately 190°K (−80°C), the constant × absolute temperature (KT) parameter of the typical devices reduces to such a level that the devices cease to operate. Three possible solutions to this difficulty have been explored:

(1) Increasing the system supply voltage and providing thermal insulation for standard integrated circuits.

(2) Neutralizing the capacity of the leads between the crystals and oscillator to allow physical separation between the crystal and its oscillator.

(3) Utilizing *n*-channel field-effect transistor (FET) semiconductor devices.

Experiments have shown that it is possible to create sufficient heat within standard insulated circuits (Fig. 1) to maintain the semiconductors above their critical lower

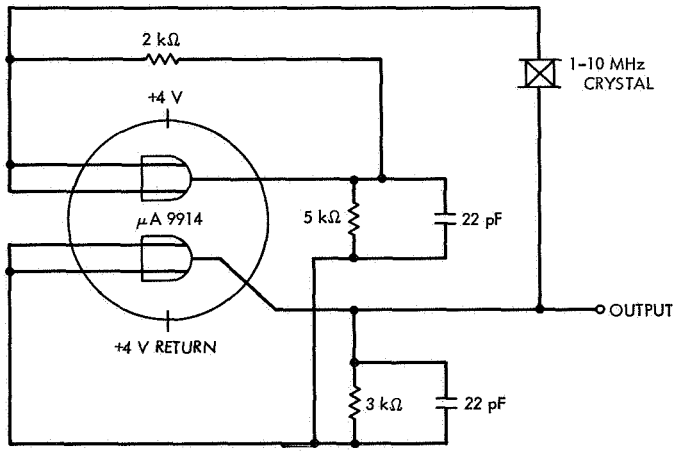


Fig. 1. Integrated circuit oscillator schematic

temperature limit. Successful operation of crystal oscillators of this type has been obtained in LN_2 with the supply voltage increased to 20 V, or more than twice the normal rating. However, the necessity of changing this supply voltage with the oscillator temperature is not a simple control problem and the crystal frequency varies with the supply voltage.

Figure 2 represents another possible solution that has been explored. Through the use of the guarded leads between the crystal and oscillator, it has been possible to separate these devices by as much as 3 m. The negative feedback signal applied to the inner shield reduces the effective center conductor capacities and prevents the oscillator from running as a multivibrator through the stray lead capacity. This technique would allow the oscillators to be located outside of the chamber in a more benign environment. Triaxial cables would be required between each oscillator and its crystal to minimize the crosstalk problem. Multiple triaxial high-vacuum feedthroughs and connectors are not readily available and would complicate the dual extreme high-vacuum seal in the Molsink.

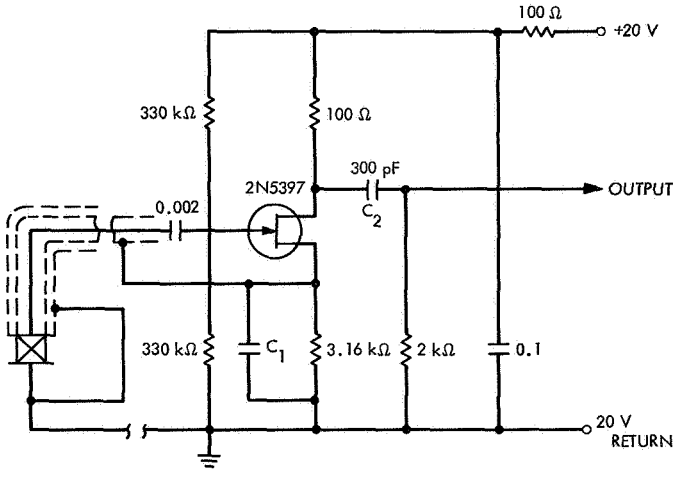
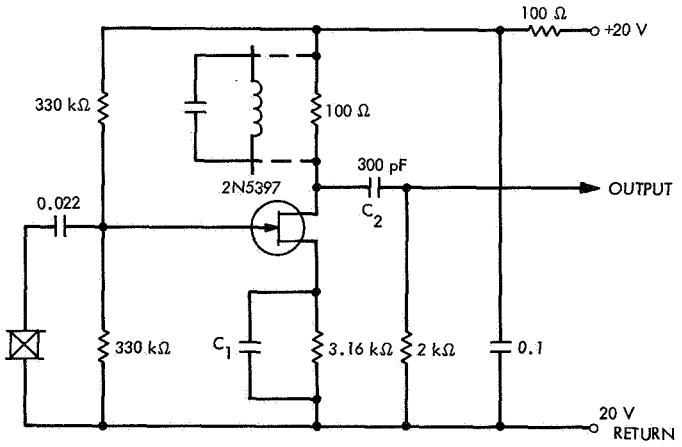


Fig. 2. Crystal to oscillator lead capacity neutralization

Since the characteristics of *n*-channel FET-type semiconductor devices do not involve the KT parameter, they will still operate at the 14°K temperature. Two types of oscillators have been utilized for microbalance measurements. The circuit diagram in Fig. 3 indicates that either a low-q tuned circuit or a 100-Ω resistor may be used in the oscillator circuit. The output from the tuned circuit is about twice the 0.5-V signal from the resistive load; however, there is less change in signal level with frequency when the resistive load is utilized. The tuned circuit has been found necessary in the 25–30 MHz frequencies. Since the q of the tuned circuit is approximately 100, a change of 0.1 MHz can introduce a change in the output level, which in turn relates to the crystal drive level. In fact, the basic uncertainty in the application of this circuit is the lack of control of the crystal drive current. The voltage sensitivity for a 10 MHz crystal is approximately 100 Hz per 5% change in the primary voltage supply. The capacity loading also has



FREQUENCY, MHz	C_1, pF	C_2, pF
5	200	300
10	56	300
20-30	22	300
3	680	0.0033

TUNED CIRCUIT MAY BE NEEDED FOR 25-30 MHz
ALL RESISTORS ARE DEPOSITED CARBON OR METAL

Fig. 3. FET oscillator circuit

a significant effect on the frequency. In spite of these basic limitations, the oscillator circuit has been found to be stable for ½ h (± 1 Hz in 10^8) at stable 14°K high-vacuum conditions.

Minimum chamber contamination from the electronic components has been achieved by fabricating the oscillator circuit as shown in Fig. 4. With the exception of resistors (the temperature coefficient of composition resistors causes an excessive change in value between 14 and 300°K), standard electronic components have been utilized. Experimentally, it has been observed that if the circuit and crystal will operate when cooled to LN_2 temperature, no further problems are encountered at the lower temperatures. However, a number of crystals have been found to fail at the LN_2 temperature and it has not been possible to establish the source of this difficulty. These same crystals have been used successfully in room temperature applications.

To date, at least a dozen different oscillators have been fabricated and tested in the 14°K environment. This process has been evolutionary in that the mounting, handling, thermal control, protection from sublimation coating, and addition of a heater lamp to clear the crystal of excessive cryo deposits have necessitated changes from oscillator to subsequent oscillator. A buffer circuit (Fig. 4) currently under investigation will, hopefully, provide isolation of the oscillator circuit from the cable loading and reduce crosstalk between the oscillators. As many as five separate oscillators have been used to attempt to measure the sublimation rate during sublimation pumping and the rate of cryo deposition. In addition to providing mass and temperature information, by utilizing proper modulation techniques, these devices can telemeter other test information out of the double-wall vacuum system thus permitting a minimum number of chamber penetrations.

References

1. Warner, A. W., "Microweighing With Resonating Quartz Plates," in *Ultra-Micro Weight Determination in Controlled Environments: Volume 12*, pp. 77–114, Plenum Press, New York, 1962.
2. Muller, R. M., and White, W., "Direct Gravimetric Calibration of a Quartz Crystal Microbalance," *Rev. Sci. Instr.*, Vol. 39, No. 3, pp. 291–295, Mar. 1968.
3. Stephens, J. B., "Space Molecular Sink Simulator Facility," paper presented at the Fourth Aerospace Mechanisms Symposium, the University of Santa Clara, Santa Clara, Calif., May 22 and 23, 1969.

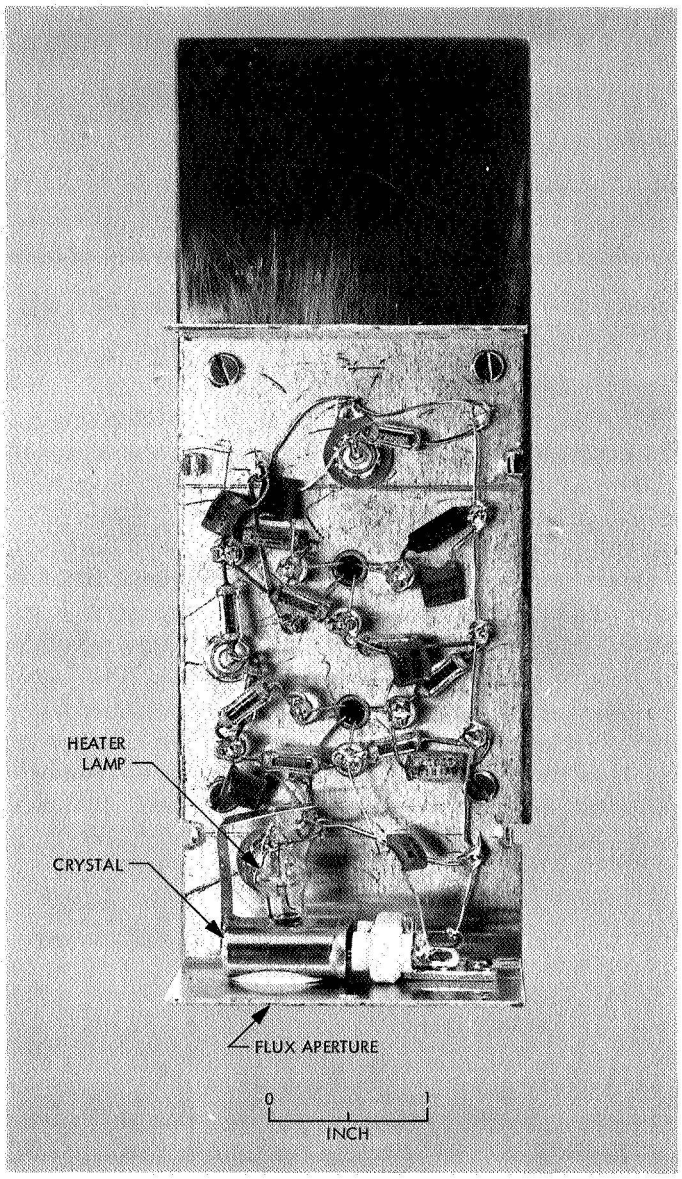


Fig. 4. FET oscillator fabrication technique

XIV. Solid Propellant Engineering

PROPULSION DIVISION

A. Heat Transfer in a Solid Propellant Grain During Ignition With Radiation Penetration and Convection Boundary Conditions, *R. L. Klaus*

1. Introduction

The proper mathematical description of many phenomena associated with the combustion of solid propellants requires a description of the mechanism of the transport of energy away from the propellant surface and into the interior of the grain. For transient phenomena such as ignition, extinction, oscillatory burning, etc., unsteady-state equations must be written and solved. The present treatment ignores the contribution due to the motion of the burning surface; therefore, it properly applies only to that part of an ignition transient during which the motion of the propellant surface is negligible. A more complete treatment of the problem to include unsteady motion of the burning surface will be forthcoming in the future.

The unsteady heat-conduction equation in solids has been discussed many times in the literature. There are, however, several disadvantages to the formulations that presently appear in the literature on solid propellant ignition and flame spreading (Refs. 1–3). First of all, they usually do not account for radiation penetration into the interior of the propellant. There is reason to believe that in many environments this can be an important effect. In fact, in modelling ignition in the arc-image furnace, this is an absolutely crucial effect and there is no doubt but that to correctly explain these experimental results for many typical present propellants one must carefully account for radiation penetration.

Secondly, many present results are formulated in terms of improper (although convergent) integrals. This is a distinct disadvantage from the standpoint of the required numerical integration.

Thirdly, the heat-conduction problem is usually solved with the boundary condition that either surface temperature or its first derivative with respect to distance into the propellant is given as a function of time. In studies of flame spreading, however, neither of these is known *a priori*. The bulk temperature of the gas and a heat-transfer coefficient are generally specified as functions of time, and the surface temperature and its first derivative are actually results of the calculation.

The present treatment circumvents all three of these problems. The possibility of radiation penetration is fully taken into account through a "heat-generation" term in the heat-conduction equation. Solutions which are written in terms of integrals are all manipulated so that all convergent integrals are proper, thus facilitating numerical calculations, and where applicable, the region of convergence of the solution is indicated. Finally, the so-called convection boundary condition is treated as one of the cases so that these results may be more directly applied to flame spreading or ignition by a convective heat-transfer stimulus.

2. Mathematical Statement and Solution of the Problem

The mathematical statement of the problem begins with the well-known one-dimensional unsteady heat-conduction

equation with a heat-generation term $G(x,t)$ written in Table 1. T is the temperature, α is thermal diffusivity (here taken to be constant), t is time, and x is distance measured into the propellant from its surface. The heat-generation term is typically used to account for heat release of sub-surface chemical reactions but is equally suited to account for radiation penetration. The radiation penetration of translucent solids may be characterized by considering that at every level x and time t a certain amount of the incident radiation is absorbed and thus appears at that position and time in exactly the same way as energy released due to sub-surface chemical reaction. It should be noted that an energy input to the system (which corresponds to radiation absorption) corresponds to a positive G as it appears in the equation mentioned above.

Table 1. General solutions of the heat conduction equation in a semi-infinite solid with heat generation

Basic equations

$$\frac{\partial^2 T}{\partial x^2} = \frac{1}{\alpha}\frac{\partial T}{\partial t} - G(x,t)$$

Boundary conditions which are always applied:	$\begin{cases} T(x,0) = F(x) \\ T(\infty,t) = 0 \end{cases}$
Additional boundary conditions:	
Case 1 (surface temperature derivative specification)	$\frac{\partial T}{\partial x}(0,t) = -H(t)$
Case 2 (surface temperature specification)	$T(0,t) = H(t)$
Case 3 (convection)	$hT(0,t) - \frac{\partial T}{\partial x}(0,t) = H(t)$

Solutions

Case 1	$T = T_{s1} + T_{i1} + T_{g1}$
Case 2	$T = T_{s2} + T_{i1} + T_{g1} + T_{i2} + T_{g2}$
Case 3	$T = T_{s3} + T_{i1} + T_{g1} + T_{i3} + T_{g3}$

Details of the functions

T_{s1} (surface temperature derivative specification)

$$T_{s1} = \frac{\partial}{\partial t}\int_0^t H(t-\lambda)\left\{2\left(\frac{\alpha\lambda}{\pi}\right)^{1/2}\exp\left(\frac{-x^2}{4\alpha\lambda}\right) - x\,\mathrm{erfc}\left[\frac{x}{2(\alpha\lambda)^{1/2}}\right]\right\}d\lambda$$

$$\frac{\partial T_{s1}}{\partial x} = -\frac{\partial}{\partial t}\int_0^t H(t-\lambda)\,\mathrm{erfc}\left[\frac{x}{2(\alpha\lambda)^{1/2}}\right]d\lambda$$

$T_{s1}(0,t) = 2\frac{\partial}{\partial t}\int_0^t H(t-\lambda)\left(\frac{\alpha\lambda}{\pi}\right)^{1/2}d\lambda$	$T_{s1}(x,0) = 0$
$\frac{\partial T_{s1}}{\partial x}(0,t) = -H(t), \qquad t>0$	$\frac{\partial T_{s1}}{\partial x}(x,0) = 0$

Table 1 (contd)

T_{i1} (initial temperature distribution)

$$T_{i1} = \frac{-1}{2(\pi\alpha t)^{1/2}} \int_0^\infty F(\xi) \left\{ \exp\left[\frac{-(x-\xi)^2}{4\alpha t}\right] + \exp\left[\frac{-(x+\xi)^2}{4\alpha t}\right] \right\} d\xi$$

$$\frac{\partial T_{i1}}{\partial x} = \frac{1}{4[\pi(\alpha t)^3]^{1/2}} \int_0^\infty F(\xi) \left\{ (x-\xi) \exp\left[\frac{-(x-\xi)^2}{4\alpha t}\right] + (x+\xi) \exp\left[\frac{-(x+\xi)^2}{4\alpha t}\right] \right\} d\xi$$

$$T_{i1} = F(x) - \frac{1}{2}\left\{ \int_0^x F'(\xi) \operatorname{erfc}\left[\frac{x-\xi}{2(\alpha t)^{1/2}}\right] d\xi - \int_x^\infty F'(\xi) \operatorname{erfc}\left[\frac{\xi-x}{2(\alpha t)^{1/2}}\right] d\xi - \int_0^\infty F'(\xi) \operatorname{erfc}\left[\frac{x+\xi}{2(\alpha t)^{1/2}}\right] d\xi \right\}$$

$$\frac{\partial T_{i1}}{\partial x} = F'(x) - F'(0) \operatorname{erfc}\left[\frac{x}{2(\alpha t)^{1/2}}\right] - \frac{1}{2}\left\{ \int_0^x F''(\xi) \operatorname{erfc}\left[\frac{x-\xi}{2(\alpha t)^{1/2}}\right] d\xi - \int_x^\infty F''(\xi) \operatorname{erfc}\left[\frac{\xi-x}{2(\alpha t)^{1/2}}\right] + \int_0^\infty F''(\xi) \operatorname{erfc}\left[\frac{x+\xi}{2(\alpha t)^{1/2}}\right] d\xi \right\}$$

$$T_{i1}(0,t) = F(0) + \int_0^\infty F'(\xi) \operatorname{erfc}\left[\frac{\xi}{2(\alpha t)^{1/2}}\right] d\xi$$

$$\frac{\partial T_{i1}}{\partial x}(0,t) = 0, \qquad t > 0$$

$$T_{i1}(x,0) = F(x)$$

$$\frac{\partial T_{i1}}{\partial x}(x,0) = F'(x)$$

T_{g1} (heat generation or radiation penetration)

$$T_{g1} = \frac{\partial}{\partial t} \int_0^t \left\{ \int_0^\infty G(\xi, t-\lambda) \left\{ \left(\frac{\alpha\lambda}{\pi}\right)^{1/2} \left[\exp\left(\frac{-(x-\xi)^2}{4\alpha\lambda}\right) + \exp\left(\frac{-(x+\xi)^2}{4\alpha\lambda}\right) \right] - \left(\frac{\xi+x}{2}\right) \operatorname{erfc}\left[\frac{\xi+x}{2(\alpha\lambda)^{1/2}}\right] \right\} d\xi \right.$$
$$\left. - \int_0^x G(\xi, t-\lambda) \left(\frac{x-\xi}{2}\right) \operatorname{erfc}\left[\frac{x-\xi}{2(\alpha\lambda)^{1/2}}\right] d\xi - \int_x^\infty G(\xi, t-\lambda) \left(\frac{\xi-x}{2}\right) \operatorname{erfc}\left[\frac{\xi-x}{2(\alpha\lambda)^{1/2}}\right] d\xi \right\} d\lambda$$

$$\frac{\partial T_{g1}}{\partial x} = -\frac{1}{2} \frac{\partial}{\partial t} \int_0^t \left\{ \int_0^\infty G(\xi, t-\lambda) \operatorname{erfc}\left[\frac{\xi+x}{2(\alpha\lambda)^{1/2}}\right] d\xi \right.$$
$$\left. + \int_0^x G(\xi, t-\lambda) \operatorname{erfc}\left[\frac{x-\xi}{2(\alpha\lambda)^{1/2}}\right] d\xi - \int_x^\infty G(\xi, t-\lambda) \operatorname{erfc}\left[\frac{\xi-x}{2(\alpha\lambda)^{1/2}}\right] d\xi \right\} d\lambda$$

$$T_{g1}(0,t) = \frac{\partial}{\partial t} \int_0^t \int_0^\infty G(\xi, t-\lambda) \left\{ 2\left(\frac{\alpha\lambda}{\pi}\right)^{1/2} \exp\left(\frac{-\xi^2}{4\alpha\lambda}\right) - \xi \operatorname{erfc}\left[\frac{\xi}{2(\alpha\lambda)^{1/2}}\right] \right\} d\xi\, d\lambda$$

$$\frac{\partial T_{g1}}{\partial x}(0,t) = 0$$

$$T_{g1}(x,0) = 0$$

$$\frac{\partial T_{g1}}{\partial x}(x,0) = 0$$

T_{s2} (surface temperature derivative specification)

$$T_{s2} = \frac{\partial}{\partial t} \int_0^t H(t-\lambda) \operatorname{erfc}\left[\frac{x}{2(\alpha\lambda)^{1/2}}\right] d\lambda$$

$$\frac{\partial T_{s2}}{\partial x} = -\frac{\partial}{\partial t} \int_0^t \frac{H(t-\lambda)}{(\pi\alpha\lambda)^{1/2}} \exp\left(\frac{-x^2}{4\alpha\lambda}\right) d\lambda$$

Table 1 (contd)

T_{s2} (contd)	
$T_{s2}(0,t) = H(t), \qquad t>0$	$T_{s2}(x,0) = 0$
$\dfrac{\partial T_{s2}}{\partial x}(0,t) = -\dfrac{\partial}{\partial t}\displaystyle\int_0^t \dfrac{H(t-\lambda)}{(\pi\alpha\lambda)^{1/2}}\,d\lambda, \qquad t>0$	$\dfrac{\partial T_{s2}}{\partial x}(x,0) = 0$
T_{i2} (initial temperature distribution)	
$T_{i2} = \dfrac{1}{(\pi\alpha t)^{1/2}}\displaystyle\int_0^\infty F(\xi)\exp\left[\dfrac{-(x+\xi)^2}{4\alpha t}\right]d\xi$	
$\dfrac{\partial T_{i2}}{\partial x} = \dfrac{-1}{2[\pi(\alpha t)^3]^{1/2}}\displaystyle\int_0^\infty F(\xi)(x+\xi)\exp\left[\dfrac{-(x+\xi)^2}{4\alpha t}\right]d\xi$	
$T_{i2} = -F(0)\operatorname{erfc}\left[\dfrac{x}{2(\alpha t)^{1/2}}\right] - \displaystyle\int_0^\infty F'(\xi)\operatorname{erfc}\left[\dfrac{x+\xi}{2(\alpha t)^{1/2}}\right]d\xi$	
$\dfrac{\partial T_{i2}}{\partial x} = \dfrac{F(0)}{(\pi\alpha t)^{1/2}}\exp\left(\dfrac{-x^2}{4\alpha t}\right) + F'(0)\operatorname{erfc}\left[\dfrac{x}{2(\alpha t)^{1/2}}\right] + \displaystyle\int_0^\infty F''(\xi)\operatorname{erfc}\left[\dfrac{x+\xi}{2(\alpha t)^{1/2}}\right]d\xi$	
$T_{i2}(0,t) = -F(0) - \displaystyle\int_0^\infty F'(\xi)\operatorname{erfc}\left[\dfrac{\xi}{2(\alpha t)^{1/2}}\right]d\xi, \qquad t>0$	$T_{i2}(x,0) = 0$
$\dfrac{\partial T_{i2}}{\partial x}(0,t) = \dfrac{F(0)}{(\pi\alpha t)^{1/2}} + F'(0) + \displaystyle\int_0^\infty F''(\xi)\operatorname{erfc}\left[\dfrac{\xi}{2(\alpha t)^{1/2}}\right]d\xi, \qquad t>0$	$\dfrac{\partial T_{i2}}{\partial x}(x,0) = 0$
T_{g2} (heat generation or radiation penetration)	
$T_{g2} = \dfrac{\partial}{\partial t}\displaystyle\int_0^t\int_0^\infty G(\xi,t-\lambda)\left\{(\xi+x)\operatorname{erfc}\left[\dfrac{\xi+x}{2(\alpha\lambda)^{1/2}}\right] - 2\left(\dfrac{\alpha\lambda}{\pi}\right)^{1/2}\exp\left[\dfrac{-(x+\xi)^2}{4\alpha\lambda}\right]\right\}d\xi\,d\lambda$	
$\dfrac{\partial T_{g2}}{\partial x} = \dfrac{\partial}{\partial t}\displaystyle\int_0^t\int_0^\infty G(\xi,t-\lambda)\operatorname{erfc}\left[\dfrac{\xi+x}{2(\alpha\lambda)^{1/2}}\right]d\xi$	
$T_{g2}(0,t) = \dfrac{\partial}{\partial t}\displaystyle\int_0^t\int_0^\infty G(\xi,t-\lambda)\left\{\xi\operatorname{erfc}\left[\dfrac{\xi}{2(\alpha\lambda)^{1/2}}\right] - 2\left(\dfrac{\alpha\lambda}{\pi}\right)^{1/2}\exp\left(\dfrac{-\xi^2}{4\alpha\lambda}\right)\right\}d\xi\,d\lambda$	$T_{g2}(x,0) = 0$
$\dfrac{\partial T_{g2}}{\partial x} = \dfrac{\partial}{\partial t}\displaystyle\int_0^t\int_0^\infty G(\xi,t-\lambda)\operatorname{erfc}\left[\dfrac{\xi}{2(\alpha\lambda)^{1/2}}\right]d\xi\,d\lambda$	$\dfrac{\partial T_{g2}}{\partial x}(x,0) = 0$
T_{s3} (convection boundary condition)	
$T_{s3} = \dfrac{1}{h}\dfrac{\partial}{\partial t}\displaystyle\int_0^t H(t-\lambda)\left\{\operatorname{erfc}\left[\dfrac{x}{2(\alpha\lambda)^{1/2}}\right] - \exp(hx+\alpha h^2\lambda)\operatorname{erfc}\left[h(\alpha\lambda)^{1/2} + \dfrac{x}{2(\alpha\lambda)^{1/2}}\right]\right\}d\lambda$	
$\dfrac{\partial T_{s3}}{\partial x} = -\dfrac{\partial}{\partial t}\displaystyle\int_0^t H(t-\lambda)\exp(hx+\alpha h^2\lambda)\operatorname{erfc}\left[h(\alpha\lambda)^{1/2} + \dfrac{x}{2(\alpha\lambda)^{1/2}}\right]d\lambda$	

Table 1 (contd)

T_{s3} (contd)	
$T_{s3}(0,t) = \frac{1}{h}\frac{\partial}{\partial t}\int_0^t H(t-\lambda)\{1-\exp(\alpha h^2\lambda)\operatorname{erfc}[h(\alpha\lambda)^{1/2}]\}\,d\lambda$	$T_{s3}(x,0) = 0$
$\frac{\partial T_{s3}}{\partial x}(0,t) = -\frac{\partial}{\partial t}\int_0^t H(t-\lambda)\exp(\alpha h^2\lambda)\operatorname{erfc}[h(\alpha\lambda)^{1/2}]\,d\lambda\,, \qquad t>0$	$\frac{\partial T_{s3}}{\partial x}(x,0) = 0$
T_{i3} (initial temperature distribution)	
$T_{i3} = h\int_0^\infty F(\xi)\exp[h(x+\xi)+\alpha h^2 t]\operatorname{erfc}\left[h(\alpha t)^{1/2}+\frac{x+\xi}{2(\alpha t)^{1/2}}\right]d\xi$ $\frac{\partial T_{i3}}{\partial x} = h\left\{T_{i3} - \frac{1}{(\pi\alpha t)^{1/2}}\int_0^\infty F(\xi)\exp\left[\frac{-(x+\xi)^2}{4\alpha t}\right]d\xi\right\}$	
$T_{i3}(0,t) = h\int_0^\infty F(\xi)\exp(h\xi+\alpha h^2 t)\operatorname{erfc}\left[h(\alpha t)^{1/2}+\frac{\xi}{2(\alpha t)^{1/2}}\right]d\xi$ $\frac{\partial T_{i3}}{\partial x}(0,t) = h\left[T_{i3}(0,t) - \frac{1}{(\pi\alpha t)^{1/2}}\int_0^\infty F(\xi)\exp\left(\frac{-\xi^2}{4\alpha t}\right)d\xi\right]$	$T_{i3}(x,0) = 0$ $\frac{\partial T_{i3}}{\partial x}(0,t) = 0$
T_{g3} (heat generation or radiation penetration)	
$T_{g3} = \frac{1}{h}\frac{\partial}{\partial t}\int_0^t\int_0^\infty G(\xi,t-\lambda)\left\{[h(x+\xi)+1]\operatorname{erfc}\left[\frac{x+\xi}{2(\alpha t)^{1/2}}\right] - 2h\left(\frac{\alpha t}{\pi}\right)^{1/2}\exp\left[\frac{-(x+\xi)^2}{4\alpha t}\right]\right.$ $\left. - \exp[h(x+\xi)+h^2\alpha t]\operatorname{erfc}\left[h(\alpha t)^{1/2}+\frac{x+\xi}{2(\alpha t)^{1/2}}\right]\right\}d\xi\,d\lambda$ $\frac{\partial T_{g3}}{\partial x} = \frac{\partial}{\partial t}\int_0^t\int_0^\infty G(\xi,t-\lambda)\left\{\operatorname{erfc}\left[\frac{x+\xi}{2(\alpha\lambda)^{1/2}}\right] - \exp[h(x+\xi)+h^2\alpha\lambda]\operatorname{erfc}\left[h(\alpha\lambda)^{1/2}+\frac{x+\xi}{2(\alpha\lambda)^{1/2}}\right]\right\}d\xi\,d\lambda$	
$T_{g3}(0,t) = \frac{1}{h}\frac{\partial}{\partial t}\int_0^t\int_0^\infty G(\xi,t-\lambda)\left\{(h\xi+1)\operatorname{erfc}\left[\frac{\xi}{2(\alpha t)^{1/2}}\right] - 2h\left(\frac{\alpha t}{\pi}\right)^{1/2}\exp\left(\frac{-\xi^2}{4\alpha t}\right)\right.$ $\left. - \exp(h\xi+h^2\alpha t)\operatorname{erfc}\left[h(\alpha t)^{1/2}+\frac{\xi}{2(\alpha t)^{1/2}}\right]\right\}d\xi\,d\lambda$ $\frac{\partial T_{g3}}{\partial x}(0,t) = \frac{\partial}{\partial t}\int_0^t\int_0^\infty G(\xi,t-\lambda)\left\{\operatorname{erfc}\left[\frac{\xi}{2(\alpha\lambda)^{1/2}}\right]\right.$ $\left. - \exp(h\xi+h^2\alpha\lambda)\operatorname{erfc}\left[h(\alpha\lambda)^{1/2}+\frac{\xi}{2(\alpha\lambda)^{1/2}}\right]\right\}d\xi\,d\lambda$	$T_{g3}(x,0) = 0$ $\frac{\partial T_{g3}}{\partial x}(x,0) = 0$
Expansion of the condensed functions	
$\frac{\partial}{\partial t}\int_0^t P(t-\lambda)\,Q(x,\lambda)\,d\lambda = P(0)\,Q(x,t) + \int_0^t \frac{\partial P}{\partial\lambda}(\lambda)\,Q(x,t-\lambda)\,d\lambda$	

Table 1 (contd)

Definition of the complementary error function
$\operatorname{erfc}(\lambda) = \frac{2}{(\pi)^{1/2}} \int_{\lambda}^{\infty} \exp(-\xi^2)\, d\xi$

Exact equations for G are beyond the scope of this present work. The time dependence of G is primarily due to variations in the strength of the source of radiation which is an input to the problem. The distance dependence, however, is due to the translucent nature of the propellant. Homogeneous solids are often treated by considering that radiation penetration falls off exponentially. In typical composite propellants, however, the translucency is not homogeneous. Relatively transparent oxidizer crystals are embedded in a matrix of relatively opaque binder. Under such circumstances, radiation penetration is probably best treated statistically. Unpublished work by the author treats radiation penetration in composite propellants in a statistical manner. Specific formulas for G useful in composite solid propellant applications are derived. For the present, however, G will be considered a known but as yet unspecified function.

The possible boundary conditions are also listed in Table 1. Of the three required, the first two shown apply in all cases. The first permits an arbitrary initial temperature distribution; the second indicates that all temperatures are to be measured above ambient and that at a sufficiently large distance into the propellant, ambient temperature is maintained.

Three possible choices are offered for the third boundary condition, depending on the case to be considered. The interpretation of the first two cases is obvious. The third, or convection boundary condition, represents an energy balance at the solid surface, assuming heat transfer by convection from the gas to the solid. This energy balance is usually first written as

$$h^*[T_g - T(0,t)] = E_s - k\frac{\partial T}{\partial x}(0,t)$$

where h^* is the heat transfer coefficient, T_g is the bulk gas temperature, E_s is any energy absorbed at the surface due to chemical reaction, and k is the thermal conductivity. This equation can be arranged so that it is in the proper form by allowing

$$h = \frac{h^*}{k}$$

$$H(t) = \frac{h^* T_g - E_s}{K}$$

Solutions to this problem for constant h are given in Table 1. They were obtained by well-known Laplace transform techniques. Formulas have been written down for both the various temperature functions, their first derivatives with respect to time, and their limits as either x or t approaches zero. Temperature functions with subscripts s and g are written in a rather condensed form. An equation at the end of Table 1 indicates how these functions may be expanded into a form more suitable for numerical calculations. The definition of the complementary error function, used frequently in the tabulation, is also given. The temperature functions with subscript i have expanded forms which are not as obvious in their derivation. Thus, both versions are written except for T_{i3} where they are the same.

Several important qualitative observations can be made concerning the tabulated functions. If a surface temperature derivative is imposed as a boundary condition, there is no discontinuity in temperature at the surface and at $t = 0$, but the derivatives of both T_{s1} and T_{i1} have discontinuities at that point, if $H(0) \neq 0$. If a surface temperature is imposed, both T_{s2} and T_{i2} have discontinuities, the derivative of T_{i2} has a discontinuity, and the derivative of T_{s2} does not converge at that point, again provided that $H(0) \neq 0$. Under the same conditions for the convection boundary condition, only the derivative of T_{s3} is found to be discontinuous. Temperature functions with subscript g and their first derivatives are always continuous.

3. Conclusions

The functions given in Table 1 may be used to determine solutions to many important heat-conduction problems in solids. Of particular note is the fact that radiation penetration and convection boundary conditions are explicitly treated and the fact that all integrals are proper

except, of course, where they diverge. In practice, it is frequently of interest to calculate both T and its first derivative at the surface with variable h. This may be accomplished in a step-wise manner by successively considering time intervals over which h is relatively constant. A solution is determined which applies up to the end of the given time interval and that solution is used as an initial temperature distribution in the equations written for the next interval which have a different value of h. This procedure can be directly programmed for a computer so that the most general heat-conduction problem of interest in solid propellant ignition may be solved.

References

1. Most, W. J., MacDonald, B. W., Stang, P. L., and Summerfield, M., *Thrust Transient Prediction and Control of Solid Rocket Engines*, Paper 68-33. Western States Section, The Combustion Institute, Oct. 29, 1968.
2. Price, E. W., and Culick, F. E. C., "Combustion of Solid Rocket Propellants" *AIAA Professional Study Series*, 1968.
3. von Elbe, G., et al., *Chemical-Kinetic and Physical Process in Composite Solid Propellant Combustion*, Report 66307, Contract NAS1-6440. Atlantic Research Corporation, Alexandria, Va., Jan. 1967.

B. Blast Wave Propagation in Air, O. K. Heiney

1. Introduction

Of primary importance for hazard determination of large solid-propellant motors is a mechanism for predicting the peak overpressure of the blast field should the entire propellant matrix undergo a detonative decomposition. The determination of this data is important in that it will provide a worst-case analysis of the blast damage potential.

The basic purpose of this effort was to predict the attenuation of the peak pressure exhibited by such a blast wave as a function of the radial distance traveled by the wave front. The basis of the engineering or heuristic approach taken was that the overpressure decay was a changing power function of the nondimensionalized radial-wave-front travel. The analysis differs from many presented previously in that a smooth transition of decay is assumed rather than an artificial polynomial fitting approach. The analysis is correlated with experimental blast-pressure attenuation measurements, and demonstrates excellent agreement. The formulation starts from an experimentally given initial point and is carried out radially from the reference. A simplistic attempt to compute the initial point from the mass and heat of explosion of the charge has had moderate success and is described. Table 2 gives the nomenclature used in the discussion.

Table 2. Nomenclature

a_1	speed of sound
C_w	charge weight
N	decay component
P_0	decibel reference pressure
P_1	undisturbed ambient pressure
P_2	shock pressure
Q	specific chemical potential or heat of explosion of solid
R	distance to shock front
r	distance to shock point
t	time
U_1	shock front velocity
u_p	particle velocity in shock front
V	initial pressurized volume
Z	scaled radial distance
γ	ratio of specific heats in gas
ρ_1	ambient gas density
ρ_2	density in shock front

2. Discussion

The literature on blast overpressure decay is voluminous. In essence, however, the analysis falls into one of two easily distinguishable categories. The first is a simple curve fitting through experimental data, and is typified by Refs. 1 and 2. The second approach consists of sophisticated mathematical gasdynamic analysis of the phenomenon. By definition, these analyses must predict the entire flow field behind the overpressure shock, rather than just the peak-pressure decay. This distinction is illustrated by Fig. 1. A traveling blast wave is illustrated at four points in time. For a similarity solution of the entire flow field at a point in time, such as those published by Sakuri (Ref. 3), Von Neumann (Ref. 4), and Brinkley (Ref. 5), a complete description of the flow field behind the pressure front at that point is necessary. To predict only the magnitude of the overpressure, it is necessary to predict only

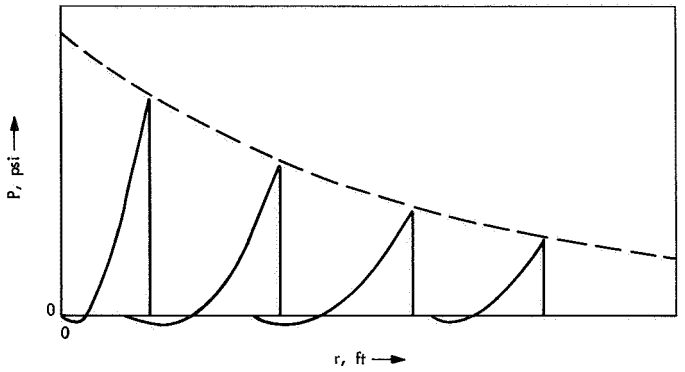


Fig. 1. Peak blast pressure vs distance from origin

the dashed line of Fig. 1. The similarity and asymptotic solutions have demonstrated very narrow ranges of applicability and whether implicitly or explicitly also contain empirical factors to fit the data to experiment.

The purpose of this effort is then to look at the middle ground between empiricism and unjustifiably elaborate analysis and attempt to predict decay of the peak overpressure demonstrated by a blast wave as it progresses radially. The central criterion of this analysis, as discussed below, is that the pressure is allowed to decay as a varying power of the distance traveled due to a change of shape of the blast bubble. A blast wave is a pressure-decaying expanding shock front which is traveling at faster than sonic velocity. The most common source of the symmetrically expanding blast waves that are considered here is from an explosion. The unclassified overpressure data which was available was for solid explosive in the forms of a cylinder, a cube, and a sphere. The weights of explosive ranged from ½ to 8 lb. This experimental data was extracted from Ref. 1 and is plotted as a solid line in Figs. 2a through 2d.

For the empirical polynomial describing the peak-pressure decay presented by Ref. 1, and repeated in Table 3, the factor $\pi = (P - P_0)/P_0$ and is thus an overpressure in atmospheres; the factor Z is a scaled radial distance where $Z = R/C_w^{1/3}$.

What is physically occurring but never specifically mentioned in the cited references is that a spherically symmetrical blast from a solid explosive can be conceptually divided into three regimes. The first would be close to the charge where the energy is filling a spherical volume by pressure increase and thermal excitation of molecules. Following this regime, the blast energy is converted to a spherically increasing thin layer at the shock front. This transition from a pressurized volume to a thin shell would occur by some mechanism similar to the piston analogy given by von Elbe (Ref. 6) for shock-wave formation, due to velocity and temperature gradients. In this regime, dissipation is due to divergence and a finite entropy increment across the shock front. Finally, the wave front can be expected to slowly change from a half-hemisphere surface to a cylindrical one.

Table 4 gives the transition points which will be used by this analysis to account for pressure decay. That is, where $N = 3$,

$$\frac{dP}{dR} \propto \frac{1}{R^3}$$

and where $N = 1.9$,

$$\frac{dP}{dR} \propto \frac{1}{R^{1.9}}$$

This distance traveled is in charge diameters, assuming a spherical charge.

Table 3. Explosive material, polynomial describing peak-pressure delay, and figure showing overpressure data

Figure	Explosive material	Polynomial
2a	8-lb cylinder of pentolite	$\pi = 10.5/Z - 136/Z^2 + 21{,}000/Z^3$
2b	4-lb cylinder of TNT	$\pi = 11.3/Z - 186/Z^2 + 19{,}210/Z^3$
2c	3¾-lb sphere of pentolite	$\pi = 8.6/Z + 295/Z^2 + 7{,}820/Z^3$
2d	½-lb cube of TNT	$\pi = 13.5/Z - 780/Z^2 + 36{,}280/Z^3$

Table 4. Transition points and decay exponents

Distance, charge diameter	Decay exponent N
<5	3.0
5–8	2.9
8–12	2.8
12–16	2.7
16–20	2.6
20–25	2.5
25–30	2.4
30–35	2.25
35–45	2.0
45–55	1.8
55–65	1.6
65–75	1.5
75–85	1.4
85–100	1.3
100–125	1.2
125–175	1.15
>175	1.00

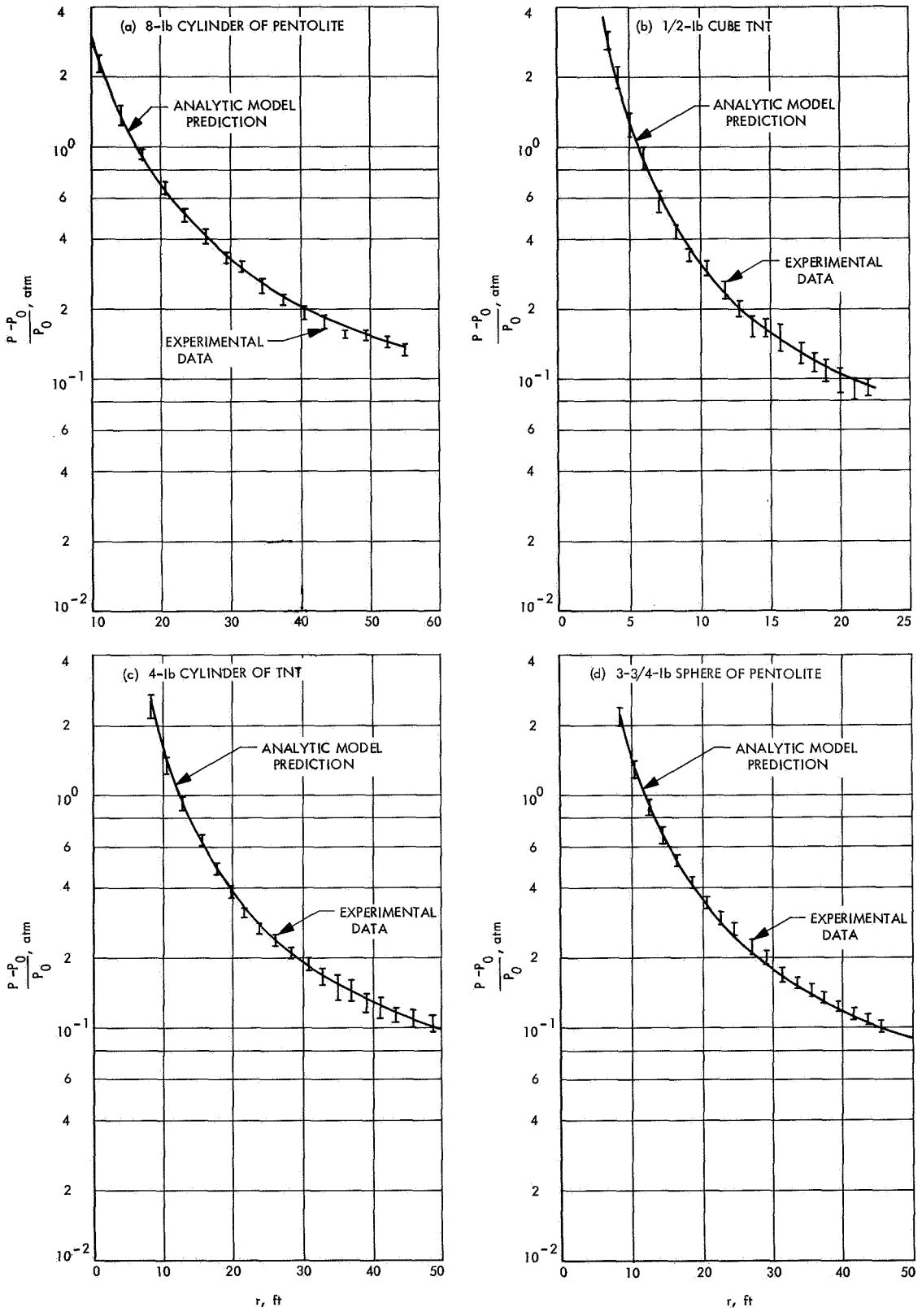


Fig. 2. Overpressure vs distance from origin

The predictions of this simplified physical approach would have the same form as the empirical curve fit methods mentioned above. This is due to the fact that close to the charge the available energy is filling a sphere and the peak pressure will decay with the cube of the distance, while at intermediate distances the pressure is spread over the surface of a hemisphere, causing a square pressure decay. This regime is in turn supplanted by an expanding cylinder and finally a constant height cylinder leading to a linear pressure decay, as would be expected for a sound wave traveling in free air (Ref. 7). The data fitting used in Ref. 1, however, leads to the physically implausible situation of a negative second term in three of the four polynomials shown in Table 3.

3. Analysis

Using this simplified model, with the assumption that the transition from a volume to a hemispherical shell to a cylindrical shell occurs in degrees as per Table 4, a pressure–travel–time history of the event may be simply described by standard gas dynamic equations presented in Ref. 8, for example,

$$\frac{dR}{dt} = U_1 \tag{1}$$

$$U_1 = a_1 \left(\frac{\gamma - 1}{2\gamma} + \frac{\gamma + 1}{2\gamma} \frac{P_2}{P_1} \right)^{1/2} \tag{2}$$

$$\frac{\rho_2}{\rho_1} = \frac{1 + \dfrac{\gamma + 1}{\gamma - 1} \dfrac{P_2}{P_1}}{\dfrac{\gamma + 1}{\gamma - 1} + \dfrac{P_2}{P_1}} \tag{3}$$

$$u_p = \frac{a_1}{\gamma} \left(\frac{P_2}{P_1} - 1 \right) \left(\frac{\dfrac{2\gamma}{\gamma + 1}}{\dfrac{P_2}{P_1} + \dfrac{\gamma - 1}{\gamma + 1}} \right)^{1/2} \tag{4}$$

$$\mathrm{dB} = 20 \log \frac{P_2 - P_1}{P_0} \tag{5}$$

With $P_0 = 2 \times 10^{-10}$ atm,

$$P_2 = \frac{R^N}{(R + \Delta R)^N} P_2' \tag{6}$$

With the knowledge of one start point comprising a distance and an overpressure measurement, a simple numerical analysis based on Eqs. (1) and (2) will predict the peak-pressure–time–distance history of the blast-wave propagation. A straightforward iterative method is possible and is modeled in the computer program. Basically, however, having an overpressure given, Eq. (2) provides a velocity which is transformed into a distance by Eq. (1). A new overpressure can then be computed with respect to the wave-shape assumptions in force at the distance the wave front is currently away from the initial blast, by means of Eq. (6). The appropriate N value for the distance is taken from Table 4. Sample output from the program that computes the decay from the initial point is listed in Table 5.

4. Results

Figures 2a and 2d show that the correlation of analysis to experiment, with the simplified theory, is excellent. This would be expected in one case since the transition points were selected to give a satisfactory fit. The utility of the approach lies in its ability to predict the decay of pressure for other charge weights, shapes, and explosives material. This is demonstrated by the range of charge weights used as comparison.

This approach relies on the ability to isolate a known starting radius and overpressure to begin the analytic procedure. In effect, an experimental boundary condition is required. An effort was made to eliminate this dependence on experiment by predicting the pressure field directly from charge weight and chemical potential. This is difficult to do in a simple, or for that matter, sophisticated analysis.

The approach assumed that the heat of explosion was transformed into pressurizing a sphere of three times the initial charge diameter and propagating outward from that point as per the decay data of Table 4. The equation used to do this was

$$P = \frac{C_w Q (\gamma - 1)}{V}$$

Table 6 gives the comparison between the experimental results and those predicted by this analysis.

Results of the analysis are not as satisfactory in this aspect. This could be expected in that no mechanism exists in the computer program to account for the effect of initial charge shape such as sphere, cylinder, tube, etc. The initial shape could be expected to have a non-negligible effect on the initial reference pressure and

Table 5. Sample output from computer program[a]

Time, s	Wavefront distance from charge, ft	Wavefront overpressure, psig	Pressure ratio, atm	Density ratio across wave	Particle velocity behind wave, ft/s	Wave velocity, ft/s	Acoustic pressure level, dB
0.0000000	8.300	32.80000	3.2312	2.20854	1042.5	1949.6	200.68
0.0001025	8.500	31.12605	3.1174	2.16119	1006.3	1918.1	200.22
0.0002068	8.700	29.57353	3.0118	2.11620	971.9	1888.5	199.78
0.0003127	8.900	28.13116	2.9136	2.07345	939.1	1860.5	199.35
0.0004202	9.100	26.78888	2.8223	2.03281	907.8	1834.1	198.92
0.0005292	9.300	25.53778	2.7372	1.99416	878.0	1809.1	198.51
0.0006398	9.500	24.36989	2.6578	1.95741	849.6	1785.5	198.10
0.0007518	9.700	23.27812	2.5835	1.92243	822.4	1763.2	197.70
0.0008652	9.900	22.25605	2.5140	1.88913	796.4	1742.0	197.32
0.0009800	10.100	21.29799	2.4488	1.85742	771.6	1721.9	196.93
0.0010962	10.300	20.39876	2.3876	1.82720	747.8	1702.8	196.56
0.0012136	10.500	19.55371	2.3301	1.79841	725.0	1684.7	196.19
0.0013323	10.700	18.75864	2.2760	1.77095	703.2	1667.4	195.83
0.0014523	10.900	18.00973	2.2251	1.74475	682.3	1651.0	195.48
0.0015734	11.100	17.30354	2.1771	1.71976	662.2	1635.4	195.13
0.0016957	11.300	16.63692	2.1317	1.69589	642.9	1620.5	194.79
0.0018191	11.500	16.00701	2.0889	1.67308	624.4	1606.4	194.46
0.0019436	11.700	15.41121	2.0483	1.65129	606.7	1592.8	194.13
0.0020692	11.900	14.84712	2.0100	1.63046	589.6	1579.9	193.80
0.0021958	12.100	14.31257	1.9736	1.61053	573.2	1567.6	193.49
0.0023233	12.300	13.80557	1.9391	1.59147	557.4	1555.8	193.17
0.0024519	12.500	13.32427	1.9064	1.57321	542.2	1544.5	192.86
0.0025814	12.700	12.86700	1.8753	1.55572	527.5	1533.7	192.56
0.0027118	12.900	12.43220	1.8457	1.53897	513.4	1523.4	192.26
0.0028431	13.100	12.01845	1.8175	1.52291	499.9	1513.5	191.97
0.0029752	13.300	11.62443	1.7907	1.50750	486.8	1504.0	191.68
0.0031082	13.500	11.24893	1.7652	1.49273	474.2	1494.9	191.40
0.0032420	13.700	10.89081	1.7408	1.47854	462.0	1486.1	191.12
0.0033765	13.900	10.54904	1.7176	1.46492	450.2	1477.8	190.84
0.0035119	14.100	10.22265	1.6954	1.45183	438.9	1469.7	190.57
0.0036480	14.300	9.91074	1.6742	1.43926	427.9	1462.0	190.30
0.0037848	14.500	9.61249	1.6539	1.42717	417.4	1454.5	190.03
0.0041285	15.000	8.98233	1.6110	1.40141	394.7	1438.7	189.44
0.0044760	15.500	8.41218	1.5722	1.37786	373.7	1424.2	188.94
0.0048271	16.000	7.89463	1.5370	1.35627	354.3	1411.0	188.32
0.0051814	16.500	7.42342	1.5049	1.33643	336.4	1398.8	187.79
0.0055389	17.000	6.99317	1.4757	1.31818	319.7	1387.6	187.27
0.0058992	17.500	6.59927	1.4489	1.30133	304.2	1377.2	186.77
0.0062622	18.000	6.23773	1.4243	1.28577	289.7	1367.6	186.28
0.0066278	18.500	5.90511	1.4017	1.27136	276.2	1358.8	185.81
0.0069958	19.000	5.59841	1.3808	1.25799	263.6	1350.6	185.34
0.0073660	19.500	5.34267	1.3634	1.24679	253.0	1343.7	184.94
0.0077381	20.000	5.10466	1.3472	1.23631	243.0	1337.2	184.55
0.0081120	20.500	4.88274	1.3321	1.22650	233.6	1331.2	184.16
0.0084876	21.000	4.67548	1.3180	1.21731	224.8	1325.5	183.78

[a]Ambient pressure = 14.7 atm, initial blast pressure = 47.5 atm, $\gamma = 1.4$, $\rho = 0.000044$, charge diameter = 0.41 ft.

apparently does. Accordingly, errors of the level of 35% exist with this approach. Further analysis should reduce this excursion, but directionality of the blast close to the change will still remain.

A final point is that the complex similarity solutions concern themselves with predicting the pressure–distance relationship for a wave at a given time. This is in contrast to the peak-pressure–travel history which this simplified approach concentrates on. These similarity solutions would have no mechanism for predicting the transition from a square to a linear decay, since an assumed form of decay is a necessary input to the momentum equation which they use. This is shown below:

Table 6. Comparison of experimental and analytic results (simplistic approach)

Charge	Distance, ft	Overpressure (experimental), psig	Overpressure (analytic), psig
8-lb cylinder of pentolite	10.5	41.3	40.2
4-lb cylinder of TNT	8.3	32.8	39.2
3¾-lb sphere of pentolite	8.0	38.3	48.0
½-lb cube of TNT	3.2	52.3	68.0

$$\frac{\rho r^{\alpha}}{\rho_0 R^{\alpha}} \frac{\partial u_p}{\partial R} + \frac{\alpha u_p}{r} = \frac{-1}{\rho a_1^2} \frac{\partial P}{\partial t}$$

with the state conditions that $\alpha = 0$ for a plane wave, $\alpha = 1$ for a cylindrical wave, and $\alpha = 2$ for a spherical wave.

References

1. Stoner, R. G., and Bleakney, W., "The Attenuation of Spherical Shock Waves in Air," *J. Appl. Phys.*, Vol. 19, July 1948.
2. *Explosive Effects and Properties*, Technical Report 65-218. U.S. Naval Ordnance Laboratory, White Oak, Maryland, June 1965.
3. Sakurai, A., "Blast Wave Theory," *Basic Developments in Fluid Dynamics*. Academic Press, Inc., New York, N.Y., 1965.
4. Von Neumann, J., and Richtmyer, R. D., *J. Appl. Phys.*, Vol. 21, p. 232, 1950.
5. Kirkwood, J. G., and Brinkley, S. R., "Theory of the Propagation of Shock Waves," *Phys. Rev.*, Vol. 51, No. 9, May 1947.
6. Von Elbe, G., *Combustion, Flames and Explosions of Gases*, Academic Press, Inc., New York, N.Y., 1961.
7. Harris, C. M., *Handbook of Noise Control*. McGraw-Hill Book Co., Inc., New York, N.Y., 1957.
8. *Equations, Tables and Charts for Compressible Flow*, NACA Report 1135. National Advisory Committee for Aeronautics, Washington, D.C., May 1953.

C. Pintle Nozzle Thrust Vector Control, *L. D. Strand*

1. Introduction

A test program was performed to investigate whether thrust vector control capabilities could be incorporated into a pintle nozzle system. The objectives of this program were to (1) determine if a pivoted pintle nozzle system is capable of producing side thrust of sufficient magnitude for motor thrust vector control, and (2) determine the criticalness of pintle nozzle position on thrust alignment for thrust magnitude control–thrust termination pintle nozzle systems.

The nozzle test system, instrumentation, procedure, and experimental results were described in SPS 37-39, Vol. IV, pp. 85–91. A pintle system was incorporated into the gas-flow nozzle test assembly (Fig. 3) of the JPL flow channel, described in SPS 37-35, Vol. IV, pp. 130–140. The pintle was capable of being canted about its pivot point at angles of ¼, ½, 1, and 1½ deg from the nozzle centerline. The nozzle surface was instrumented with static pressure orifices, allowing the pressure data to be integrated over the nozzle surface to obtain an estimate of the side forces acting on the nozzle. An example of the wall static pressure profiles obtained along the plane of the cant for a cant angle of 1 deg is shown in Fig. 4. The pressure profiles are as expected, except for the pressure perturbation located at an axial distance of approximately 3.6 in. from the nozzle flange (expansion ratio ϵ of approximately 2.4). This perturbation occurred at all cant angles, increasing in strength with increasing cant angle. The perturbation was postulated in SPS 37-39, Vol. IV, to be a recompression (shock) wave emanating from the pintle tip.

The integrated side-force data were presented in SPS 37-45, Vol. IV, pp. 83–85. The results were quite unexpected. For the smaller cant angles (¼, ½, and 1 deg),

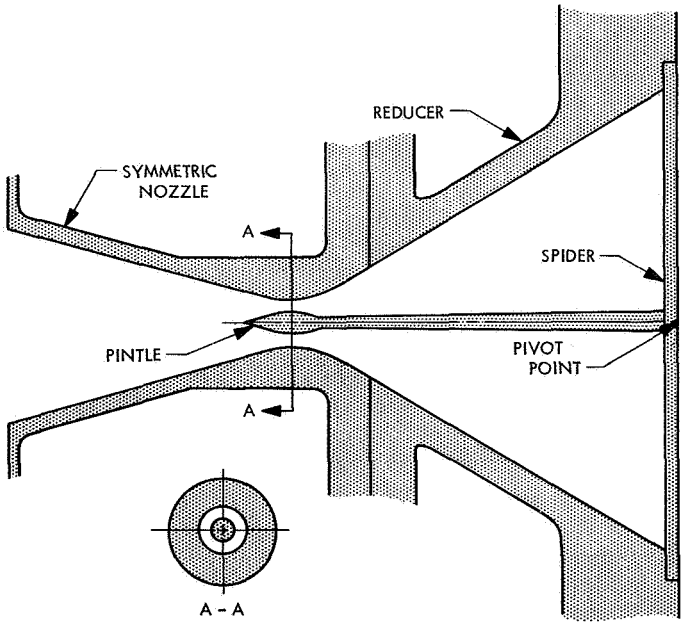


Fig. 3. Gas-flow nozzle

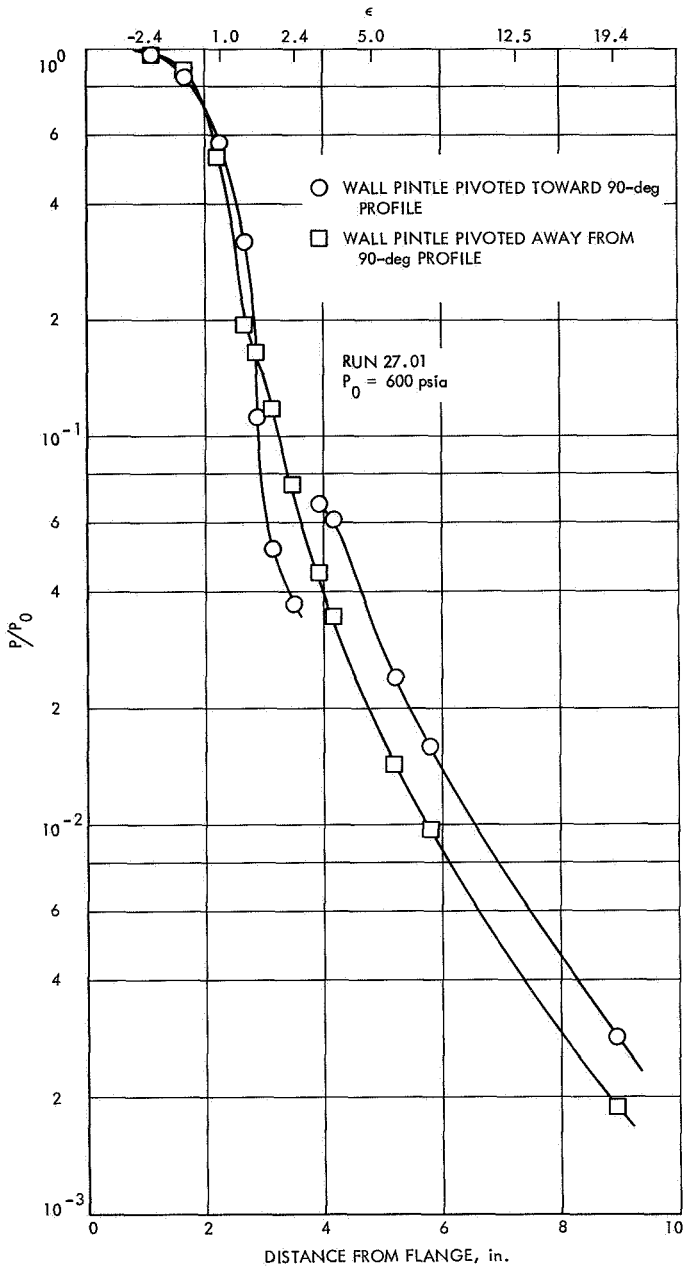


Fig. 4. Static pressure ratio versus nozzle axial distance (cant angle = 1 deg)

the expected nozzle forces resulting from the pintle asymmetry were essentially nullified by the shock wave incident on the nozzle surface. Increasing the cant angle to 1½ deg resulted in a markedly increased side force from the incident shock.

2. Two-Dimensional Nozzle Method-of-Characteristics Analysis

The nature of the wall pressure perturbation was investigated further, with the goal of understanding and possibly controlling the phenomenon. Table 7 lists the approximate locations from the nozzle flange and the accompanying ϵ values of the measured surface pressure perturbation for the axisymmetric pintle and the three largest cant angles. The location of these points is shown on a cross section of the test nozzle in Fig. 5, along with the centerline and upper-pintle contour positions for the four test cases.

A method-of-characteristics program (Ref. 1) in operation at JPL was used to calculate the wall pressure profiles

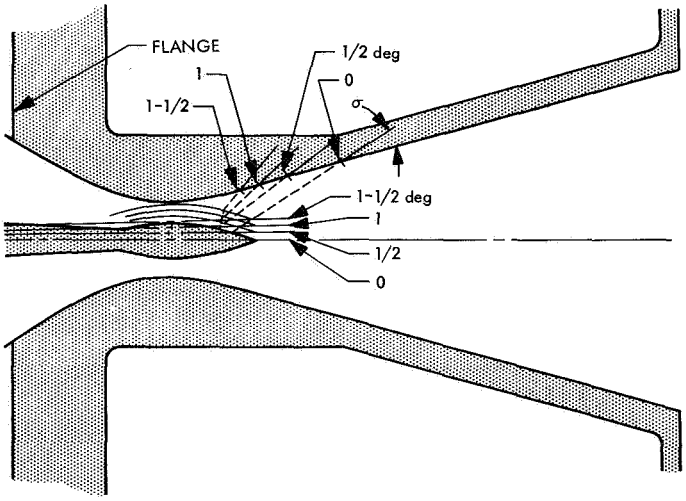


Fig. 5. Pintle nozzle cross section (0-, ½-, 1-, and 1½-deg pintle positions)

Table 7. Pressure perturbation data

Pintle cant angle, deg	Estimated distance from flange of shock incidence on nozzle surface, in.	ϵ	At incident point				
			P_1/P_0	$(P_2/P_0)/(P_1/P_0)$	M_1 one-dimensional isentropic flow table, $k = 1.4$	M_x normal shock table, $k = 1.4$	Shock angle σ, deg
0	4.8	5.0	0.020	1.02	3.20	1.01	18.2
½	4.0	3.5	0.032	1.16	2.90	1.07	21.6
1	3.6	2.4	0.036	1.89	2.82	1.33	28
1½	3.3	1.8	0.035	2.86	2.83	1.61	34.7

for the two-dimensional nozzle contour shown in Fig. 6, essentially the cross-sectional profile for the 1½-deg cant angle experimental case. The straight-line starting mach lines ($M = 1.05$) are also shown. Separate calculations were run for the upper and lower halves of the nozzle. The flow boundary aft of the pintle tip was assumed to be a straight line parallel to the nozzle centerline, as shown. The program calculated the shock waves that would be required to turn the flow aft of the pintle tip. The computed oblique shock waves are shown as dashed lines in Fig. 6. The origin of the abscissa is the pivot point of the pintle. Even taking into account the differences in two- and three-dimensional flow, there is a large discrepancy between the shock incident points of Fig. 6 and the experimental 1½-deg cant angle incident point of Fig. 5. Also, while the shock emanating from the pintle tip should be of a 360-deg nature, the test results showed, except for the largest cant angle, a marked surface pressure perturbation only in the two 90-deg quadrants adjacent to the direction of the pintle cant, i.e., in one-half of the nozzle circumference.

3. Semigraphical Shock-Wave Analysis

From the measured conditions at the nozzle wall, a semigraphical analysis was performed to attempt to determine the point of initiation of the incident shock waves. The static pressure ratio upstream of the incident shock (P_1/P_0) and the pressure ratio across the shock $(P_2/P_0)/(P_1/P_0)$ were estimated from the experimental data for each of the cant angles (Table 7). The approach

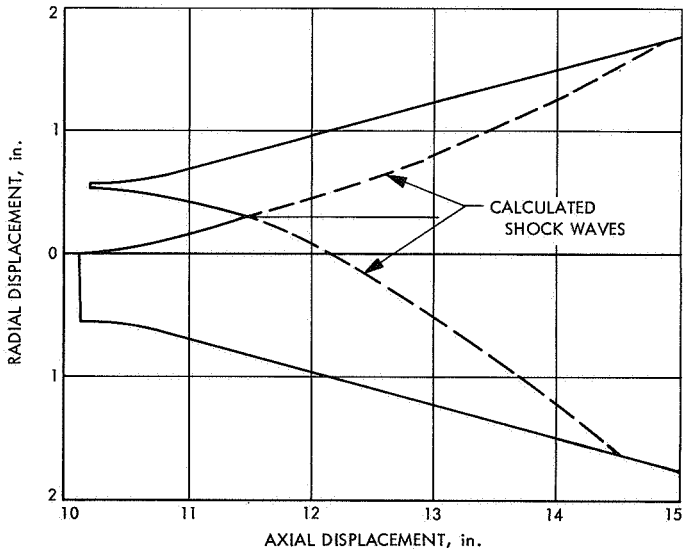


Fig. 6. Two-dimensional pintle nozzle contour (cant angle = 1½ deg)

Mach number M_1 was obtained from the one-dimensional, isentropic flow table, $k = 1.4$. From Ref. 2, "an oblique shock with an approach Mach number M_1 may, by a transformation of the coordinate system, be reduced to a normal shock with an approach Mach number $M_1 \sin\sigma$," where σ is the angle of shock to incoming flow. Therefore, for each shock pressure ratio, an approach Mach number M_x was obtained from the normal shock table, $k=1.4$. Using the coordinate system transformation, $M_x = M_1 \sin\sigma$, the shock angle σ was calculated for each cant angle (Table 7).

Assuming that the direction of the incoming flow at the nozzle surface is parallel to the surface, shock waves were drawn through each incident point at the corresponding angle from the nozzle wall and extended to the pintle surface (dashed lines in Fig. 5). The resulting points of intersection fell slightly aft of the tangency point between the curved and conical segments of the pintle contour, approaching the tangency point with increasing cant angle. The tangency points are indicated by tick marks in Fig. 5.

A possible explanation for this phenomenon has been predicted theoretically (Refs. 3 and 4) and noted experimentally (Refs. 5 and 6). Using careful instrumentation, a rise in the wall static pressure just downstream of the tangency between the circular-arc throat and the conical expansion section has been measured for nozzles with relatively small throat radius of curvature. The pressure rise is explained as a compression wave required to straighten the strong angular flow motion induced by the small throat radius of curvature.

The throat radius of curvature of the pintle and nozzle used in this test are large enough so that the recompression waves are probably ordinarily too weak to be readily detected by the relatively widely spaced static pressure orifices used in this program. If the preceding explanation is correct, then canting the pintle towards the nozzle wall in some way resulted in a detectable increase in the strength of the recompression wave at the pintle tangency region nearest the nozzle wall.

For the 1½-deg cant angle experiment, a compression wave was also measured at the nozzle wall opposite from the direction of the pintle cant (Fig. 14, p. 91, SPS 37-39, Vol. IV). The approximate point of incidence is upstream of the pintle and nozzle tangency points. This wave seems to have the characteristics of an over-expansion shock (high rate of expansion terminated by strong recompression) and is concluded to be of different origin than the previously discussed compression waves.

4. Conclusions

The nozzle side-force results calculated from the test program data were found to be strongly influenced by abrupt compression waves measured at the nozzle surface. A very rough estimate of the nozzle side forces (calculated as percentage of the nozzle theoretical, one-dimensional, isentropic axial thrust) that would have resulted if the incident shock had been absent yielded the following:

Cant angle, deg	$\Delta F/T$, %
¼	¾
½	2½
1	5
1½	5½

It is concluded that significant, predictable, and controllable side forces could be produced by an asymmetric pintle nozzle system, such as that tested, if one could, possibly through proper design of the pintle contour, eliminate the shock waves encountered in these experiments.

The second question to be answered was the sensitivity of thrust magnitude control–thrust termination pintle nozzle systems to pintle-misalignment-induced side forces. If these test results are valid, then resultant side forces should be able to be controlled to an adequately negligible level by holding to reasonable manufacturing and assembly alignment tolerances.

References

1. Prozan, R. J., *Development of a Method of Characteristics Solution for Supersonic Flow of an Ideal, Frozen, or Equilibrium Reacting Gas Mixture*, Report A 782535, Contract NAS 8-20082. Lockheed Missiles and Space Company, Research and Engineering Center, Huntsville, Ala., Apr. 1966.
2. Shapiro, A. H., *The Dynamics and Thermodynamics of Compressible Fluid Flow*, Vol. I, p. 537. The Ronald Press Co., New York, N.Y., 1953.
3. Darwell, H. M., and Badham, H., "Shock Formation in Conical Nozzles," *AIAA J.*, Vol. 1, pp. 1932–1934, 1963.
4. Migdal, D., and Kosson, R., "Shock Predictions in Conical Nozzles," *AIAA J.*, Vol. 3, pp. 1554–1556, 1965.
5. Back, L. H., and Cuffel, R. F., "Detection of Oblique Shocks in a Conical Nozzle with a Circular-Arc Throat," *AIAA J.*, Vol. 4, pp. 2219–2221, 1966.
6. Back, L. H., Massier, P. F., and Cuffel, R. F., "Flow Phenomena and Convective Heat Transfer in a Conical Supersonic Nozzle," *J. Spacecraft Rockets*, Vol. 4, pp. 1040–1047, 1967.

XV. Polymer Research

PROPULSION DIVISION

A. Hookean, Newtonian, and Maxwellian Fillers in Hookean and Standard-Solid Matrices,

J. Glucklich

1. Introduction

The effect of rigid fillers on the stress–strain behavior of a rubbery matrix has been investigated in studies conducted during the past three years (SPS 37-32, Vol. IV, pp. 120–125, SPS 37-35, Vol. IV, pp. 93–98, and SPS 37-41, Vol. IV, pp. 97–106). The composite system used was based on styrene–butadiene rubber filled with varying amounts of glass beads. It is now proposed to extend the study to fillers of different rheological behavior, viz., liquids of Newtonian and Maxwellian character. In the prior studies with rigid fillers, the change in modulus with filler content was shown to be in good agreement with the results of existing equations for E_c/E_m (Table 1), all of which are based on the Einstein viscosity equation. However, no corresponding equations exist for liquid fillers in an elastic matrix. Furthermore, if a time factor is introduced through the viscosity of the filler, the behavior of the rubber matrix can no longer be regarded as time-independent elastic. Consequently, it was considered useful to derive equations for E_c/E_m for systems with liquid (Newtonian and Maxwellian) fillers in elastic and viscoelastic (standard-solid) matrices to serve as a check to the experimental work. The rigorous approach used by Einstein and his followers in their relatively simple systems will offer almost insurmountable mathematical difficulties in the present case. Therefore, a very simplified model was sought which, although not accurately representative of the system, will nevertheless represent it to a fair degree and will offer no great mathematical difficulties. Since the case in question concerns the Young's modulus ratios only, it was decided to assume each filler particle to be surrounded by a cube of the matrix material and the behavior of each such composite element to be representative of the entire mass.[1] It was further assumed that a plane cross-section of the cube (including the cross-section through the particle) originally normal to the axis of applied force remains plane and normal under the action of load. Such a model, used by B. Paul (Ref. 1) for elastic fillers in an elastic matrix, yielded good agreement with experiments in the alloying of materials within a metal.

[1]This assumption limits the distribution of filler particles within the mass to a cubic order, whereas, in practice, a random distribution should be expected. It can be shown that for the latter case, linear and area concentrations are equal to volume concentration. This requires the power of ϕ in the formulas that follow to be modified.

Table 1. Nomenclature

A	area
E	Young's modulus
F	uniaxial force
m	E_f/E_m = modulus ratio
t	time
V	volume
ϵ	uniaxial strain
λ	uniaxial coefficient of viscosity ($\approx 3\eta$)
σ	uniaxial stress
τ_{ret}	λ/E (Kelvin) = retardation time
τ_{rel}	λ/E (Maxwell) = relaxation time
ϕ	V_f/V_c = volume concentration of fillers
Subscripts	
c	relating to composite
m	relating to matrix
f	relating to filler
$m+f$	relating to matrix and filler in parallel

In brief, the assumptions made in this analysis are as follows:

(1) An elementary cube of the matrix material containing a filler particle in its center represents the macroscopic behavior of the composite mass.

(2) The filler particles bond completely to the matrix.

(3) No separation occurs between the phases, even in the case of differing Poissons' ratio values.

(4) Plane cross-sections normal to the applied force remain plane and normal.

(5) There is no interaction between particles.

(6) The particle's shape is not specified, although the derivation is strictly correct for cubes only. For other cases, a shape factor should be included.

(7) The analysis is for uniaxial stress only.

(8) The models shown are regarded as mechanical and not rheological.

2. Hookean Filler in a Hookean Matrix

When, for a Hookean filler in a Hookean matrix (Fig. 1), an axial force F_c is applied to an elementary cube of the composite with volume V_c and cross-sectional area $V_c^{2/3}$, the axial stress σ is

$$\sigma_c = \frac{F_c}{V_c^{2/3}} \tag{1}$$

We shall consider the resulting length change of the cube to be the sum of the length changes of the pure matrix at the two ends and of the matrix plus filler at the center. Hence, the axial strain ϵ is

$$\epsilon_c = \frac{1}{V_c^{1/3}}\left[\epsilon_m\left(V_c^{1/3} - V_f^{1/3}\right) + \epsilon_{m+f} V_f^{1/3}\right] \tag{2}$$

where $V_c^{1/3}$ and $V_f^{1/3}$ are the linear dimensions of the composite and the filler, respectively. We shall also consider that the force at the cross-section through the filler is the sum of the forces on the filler and on the matrix, whence ϵ_m and ϵ_{m+f} are given by:

$$\epsilon_m = \frac{F_c}{V_c^{2/3} E_m} \tag{3}$$

and

$$\epsilon_{m+f} = \frac{F_c}{E_m\left(V_c^{2/3} - V_f^{2/3}\right) + E_f V_f^{2/3}} \tag{4}$$

Substituting Eqs. (3) and (4) for ϵ_m and ϵ_{m+f} in Eq. (2) and then using Eqs. (1) and (2) in $1/E_c = \epsilon_c/\sigma_c$, we have

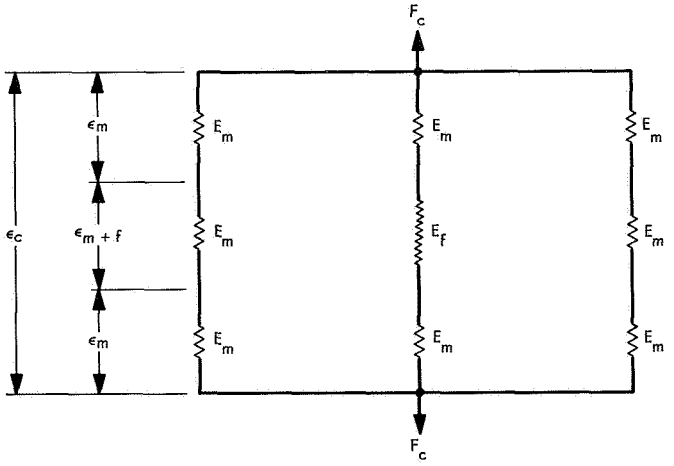


Fig. 1. Hookean filler in a Hookean matrix

$$\frac{1}{E_c} = \frac{V_c^{1/3}\left(V_c^{1/3} - V_f^{1/3}\right)}{E_m V_c^{2/3}} + \frac{V_c^{1/3} V_f^{1/3}}{E_m\left(V_c^{2/3} - V_f^{2/3}\right) + E_f V_f^{2/3}} \tag{5}$$

Noting that $V_f/V_c = \phi$ and $E_f/E_m = m$, we write Eq. (5) in the form

$$\frac{E_m}{E_c} = 1 - \phi^{1/3} + \frac{1}{\phi^{-1/3} - \phi^{1/3} + m\,\phi^{1/3}} \tag{6}$$

Equation (6) agrees with Paul's results.

When the fillers are rigid particles, $m = \infty$ and

$$\left.\frac{E_m}{E_c}\right|_{m=\infty} = 1 - \phi^{1/3} \tag{7}$$

Using a different reasoning, F. Bueche (Ref. 2) reached the same result.

When the fillers are voids, $m = 0$ and

$$\left.\frac{E_m}{E_c}\right|_{m=0} = 1 - \phi^{1/3} + \frac{1}{\phi^{-1/3} - \phi^{1/3}} \tag{8}$$

3. Newtonian Filler in a Hookean Matrix

For a Newtonian filler in a Hookean matrix (Fig. 2), Eqs. (1), (2) and (3) still apply, and for ϵ_{m+f}, we again assume the force on the central cross-section to be the sum of the forces on the filler and on the matrix. However, the force on the filler now depends on the rate of straining, as follows

$$F_c = E_m\,\epsilon_{m+f}\,A_m + \lambda_f\,\dot{\epsilon}_{m+f}\,A_f \tag{9}$$

The general solution of this differential equation is

$$\epsilon_{m+f} = \exp\left(-\frac{E_m A_m}{\lambda_f A_f}\,t\right) \times\left[\epsilon^0_{m+f} + \frac{1}{\lambda_f A_f}\int_0^t F \exp\left(\frac{E_m A_m}{\lambda_f A_f}\,t\right) dt\right] \tag{10}$$

where ϵ^0_{m+f} is the strain at $t = 0$.

For F_c = constant and $\epsilon^0_{m+f} = 0$, Eq. (10) reduces to

$$\epsilon_{m+f} = \frac{F_c}{E_m\left(V_c^{2/3} - V_f^{2/3}\right)}\left[1 - \exp\left(-\frac{V_c^{2/3} - V_f^{2/3}}{V_f^{2/3}}\,\frac{t}{\tau_{\text{ret}}}\right)\right] \tag{11}$$

where areas were expressed in terms of volumes and τ_{ret} was substituted for λ_f/E_m.

Again, as in the case of an elastic filler, Eqs. (3) and (11) are substituted for ϵ_m and ϵ_{m+f} in Eq. (2). Equations (1) and (2) are then used in $1/E_c = \epsilon_c/\sigma_c$ to give

$$\frac{1}{E_c} = \frac{V_c^{1/3}\left(V_c^{1/3} - V_f^{1/3}\right)}{E_m V_c^{2/3}} + \frac{V_c^{1/3} V_f^{1/3}}{E_m\left(V_c^{2/3} - V_f^{2/3}\right)} \times\left[1 - \exp\left(-\frac{V_c^{2/3} - V_f^{2/3}}{V_f^{2/3}}\,\frac{t}{\tau_{\text{ret}}}\right)\right] \tag{12}$$

or

$$\frac{E_m}{E_c} = 1 - \phi^{1/3} + \frac{1}{\phi^{-1/3} - \phi^{1/3}}\left[1 - \exp\left(-\left(\phi^{-2/3} - 1\right)\frac{t}{\tau_{\text{ret}}}\right)\right] \tag{13}$$

Since a Newtonian liquid acts as a rigid solid for infinite velocity of straining, putting $t = 0$ in Eq. (13) reduces it to the case of rigid particles (Eq. 7). Similarly, at $t = \infty$, Eq. (13) reduces to the case of voids (Eq. 8).

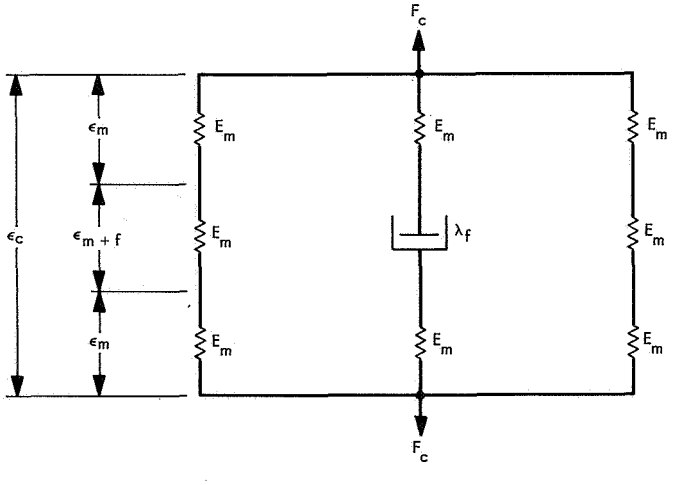


Fig. 2. Newtonian filler in a Hookean matrix

4. Maxwellian Filler in a Hookean Matrix

Here, the dashpot of Fig. 2 is replaced by a spring and dashpot in series and Eqs. (1), (2), and (3) hold as before. For ϵ_{m+f}, the reasoning used in the case of a Newtonian filler leads to the following solution:

$$\epsilon_{m+f} = \frac{F_c}{E_m\left(V_c^{2/3} - V_f^{2/3}\right) + E_f V_f^{2/3}} \left\{1 + \frac{E_f V_f^{2/3}}{E_m\left(V_c^{2/3} - V_f^{2/3}\right)} \left[1 - \exp\left(-\frac{E_m\left(V_c^{2/3} - V_f^{2/3}\right) E_f V_f^{2/3}\, t}{\left[E_m\left(V_c^{2/3} - V_f^{2/3}\right) + E_f V_f^{2/3}\right] \lambda_f}\right)\right]\right\} \quad (14)$$

From this solution, the following modulus ratio may be derived:

$$\frac{E_m}{E_c} = 1 - \phi^{1/3} + \frac{1}{\phi^{-1/3} - \phi^{1/3} + m\,\phi^{1/3}} \left\{1 + \frac{m}{\phi^{-2/3} - 1} \left[1 - \exp\left(-\frac{\left[\phi^{2/3} - 1\right] m}{\left[\phi^{-2/3} - 1\right] + m} \frac{t}{\tau_{\text{ret}}}\right)\right]\right\} \quad (15)$$

At $t = 0$, the filler should reduce to a Hookean element (Eq. 15 becomes equivalent to Eq. 6), and, similarly, at $t = \infty$, the filler should reduce to voids (Eq. 15 reduces to Eq. 8).

5. Hookean Filler in a Standard-Solid Matrix

For a Hookean filler in a standard-solid matrix (Fig. 3), the effective modulus of the matrix is time-dependent.

Since

$$\epsilon_m = \sigma_m \left\{\frac{1}{E_{m_1}} + \frac{1}{E_{m_2}} \left[1 - \exp\left(-\frac{E_{m_2}}{\lambda} t\right)\right]\right\} \quad (16)$$

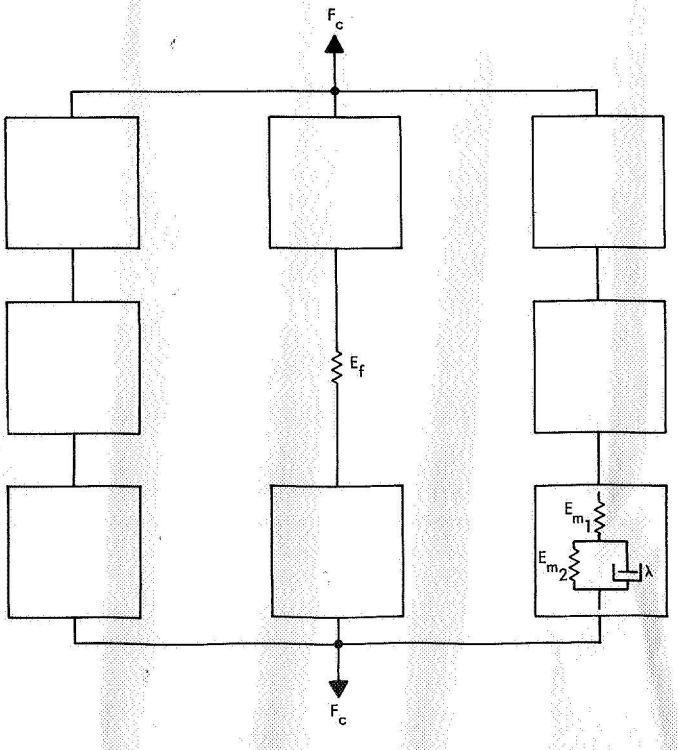


Fig. 3. Hookean filler in a standard solid matrix

then

$$E_m(t) = \left\{\frac{1}{E_{m_1}} + \frac{1}{E_{m_2}} \left[1 - \exp\left(-\frac{E_{m_2}}{\lambda} t\right)\right]\right\}^{-1} \quad (17)$$

Simple consideration will show that this system is equivalent to the one treated in *Subsection 2*, viz., of a Hookean matrix, with a time-dependent modulus replacing the constant modulus. Hence, again

$$\frac{E_m(t)}{E_c(t)} = 1 - \phi^{1/3} + \frac{1}{\phi^{-1/3} - \phi^{1/3} + m\,\phi^{1/3}} \quad (18)$$

where both E_m and E_c are time-dependent. The ratio of these two values is also time-dependent since $m = E_f/E_m(t)$ (the ratio of *effective* moduli) varies with time in accordance with

$$m = E_f \left\{\frac{1}{E_{m_1}} + \frac{1}{E_{m_2}} \left[1 - \exp\left(-\frac{t}{\tau_{\text{ret}}}\right)\right]\right\} \quad (19)$$

where

$$\tau_{\text{ret}} = \frac{\lambda}{E_{m_2}}$$

6. Newtonian Filler in a Standard Solid Matrix

Here, the spring of Fig. 3 is replaced by a dashpot with viscosity λ_f (that of the standard solid being λ_m). Equations (1) and (2) still apply, but now ϵ_m becomes

$$\epsilon_m = \frac{F_c}{V_c^{2/3}} \left\{\frac{1}{E_{m_1}} + \frac{1}{E_{m_2}} \left[1 - \exp\left(-\frac{E_{m_2}}{\lambda_m} t\right)\right]\right\} \quad (20)$$

For ϵ_{m+f}, the reasoning of *Subsection 3* is followed again

$$F_c = \dot{\epsilon}_{m+f}\ A_f\ \lambda_f + \epsilon_{m+f}\ A_m\ E_m(t) \quad (21)$$

with $E_m(t)$ as in Eq. (17).

The complete solution of the differential Eq. (21) with the boundary conditions of $\epsilon^0_{m+f} = 0$ and F_c = constant is

$$\epsilon_{m+f} = \frac{F_c}{A_m}\left\{\frac{1}{E_{m_1}} + \frac{1}{E_{m_2}}\left[1 - \exp\left(-\frac{E_{m_2}}{\lambda_m}t\right)\right]\right\} \times \left\{1 - \exp\left[-\frac{A_m t}{A_f \lambda_f\left\{\frac{1}{E_{m_1}} + \frac{1}{E_{m_2}}\left[1 - \exp\left(-\frac{E_{m_2}}{\lambda_m}\right)\right]\right\}}\right]\right\} \tag{22}$$

Following the by now standard procedure of substituting Eqs. (20) and (22) into Eq. (2), solving for $1/E_c = \epsilon_c/\sigma_c$ using Eqs. (1) and (2), and replacing V_f/V_c with ϕ, we arrive at

$$\frac{E_m(t)}{E_c(t)} = 1 - \phi^{1/3} + \frac{1 - \exp\left[-(\phi^{-2/3} - 1)\frac{t}{\tau_{\text{ret }(c)}}\right]}{\phi^{-1/3} - \phi^{1/3}} \tag{23}$$

where $E_m(t)$ is equal to the expression contained in the first brace on the right of Eq. (22),

$$\tau_{\text{ret }(c)} = \frac{\lambda_f}{E_m(t)} = \lambda_f\left\{\frac{1}{E_{m_1}} + \frac{1}{E_{m_2}}\left[1 - \exp\left(-\frac{t}{\tau_{\text{ret }(m)}}\right)\right]\right\} \tag{24}$$

is the composite retardation time (time-dependent), and

$$\tau_{\text{ret }(m)} = \frac{\lambda_m}{E_{m_2}}$$

is the matrix retardation time (constant).

Obviously, the ratio $E_m(t)/E_c(t)$ in Eq. (23) is time-dependent and, checking its value for $t = 0$ and $t = \infty$, shows agreement with Eq. (18) when $m = \infty$ and $m = 0$, respectively.

References

1. Paul, B., "Prediction of Elastic Constants of Multiphase Materials," *Trans. Met. Soc. AIME*, Vol. 218, pp. 36–41, Feb. 1960.
2. Bueche, F., "Molecular Basis for the Mullins Effect," *J. Appl. Polym. Sci.*, Vol. IV, No. 10, pp. 107–114, 1960.

B. Glassy-State Properties of Polymers: Epoxy Resins, *E. F. Cuddihy and J. Moacanin*

1. Introduction

Dynamic mechanical properties for five epoxy resins having well-defined chemical structures were previously described in SPS 37-49, Vol. III, pp. 177–183. Current theories hold that transitions in polymers in the glassy state are related to the mobility of small segments along the polymer backbone. With the aid of Hirschefelder Molecular Models, mobile segments were identified in each of the five epoxy systems and correlated with the observed dynamic transitions.

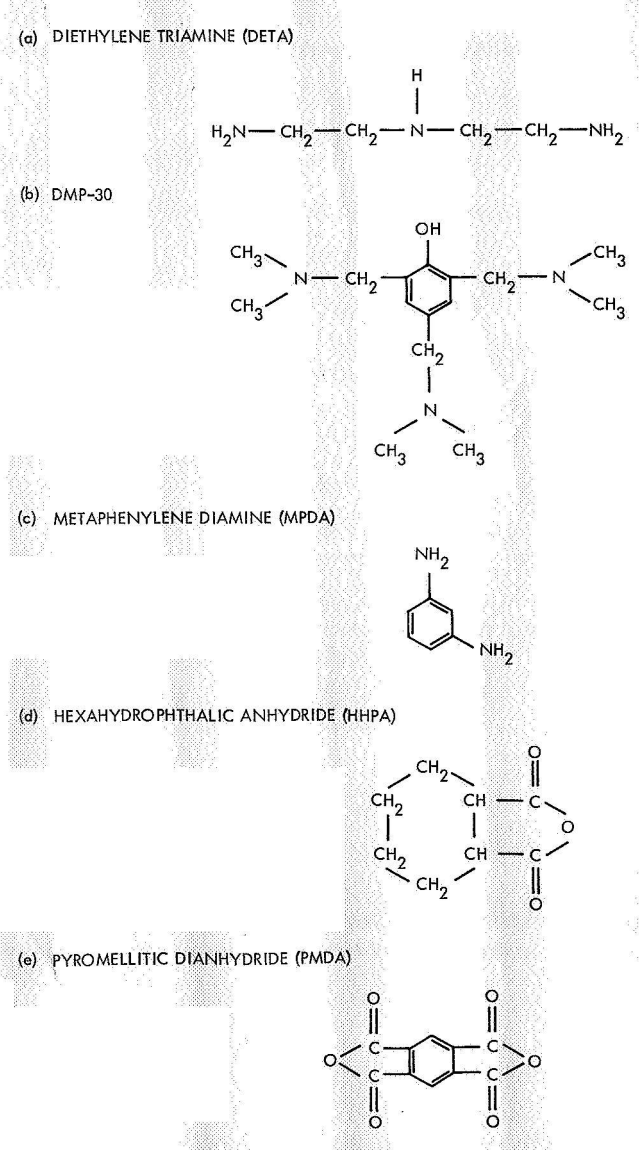


Fig. 4. Chemical structures of epoxy curatives

As each of the five epoxy resins have a common β transition, assignment of diglycidyl ether of bisphenol A (DGEBA), the common constituent, as the cause of the transition seemed straightforward. As previously demonstrated (SPS 37-49, Vol. III), it was this assignment that brought uniformity to the sparse literature on the dynamic properties of epoxies. However, the failure to find transitions for the diethylene triamine (DETA) and hexahydrophthalic anhydride (HHPA) cure, and an attempt to achieve a deeper understanding of the DGEBA role in the β transition, prompted the present analysis.

2. Chemical Structures

Five DGEBA-based epoxy resins were prepared using different curing agents (Fig. 4). Examination of molecular models disclosed that DGEBA and two of the curatives contain mobile moieties (Fig. 5).

The mobile segment in DGEBA consists of the dimethylene ether linkage of bisphenol A (Fig. 5a). As for the curatives, one mobile segment was found for DETA (Fig. 5b) while another was found for HHPA (Fig. 5c). Thus, all the five epoxy resins had in common the DGEBA-introduced moiety. In addition, two of the epoxies (DETA and HHPA) contained a single curative-introduced group each, whereas three of the epoxies [DMP-30, metaphenylene diamine (MPDA), and pyromellitic dianhydride (PMDA)] contained none.

3. Dynamic Mechanical Properties

Figures 6a through 6e show the plots of loss tangent (tan θ) for each of the epoxies versus reciprocal temperature in degrees Kelvin, where the temperature range covered extends from $-196°C$ to just short of the epoxy glass transition temperature.

All five tan θ plots are similar, each having a single large transition (hereafter referred to as the β transition) followed with decreasing temperature by an apparent

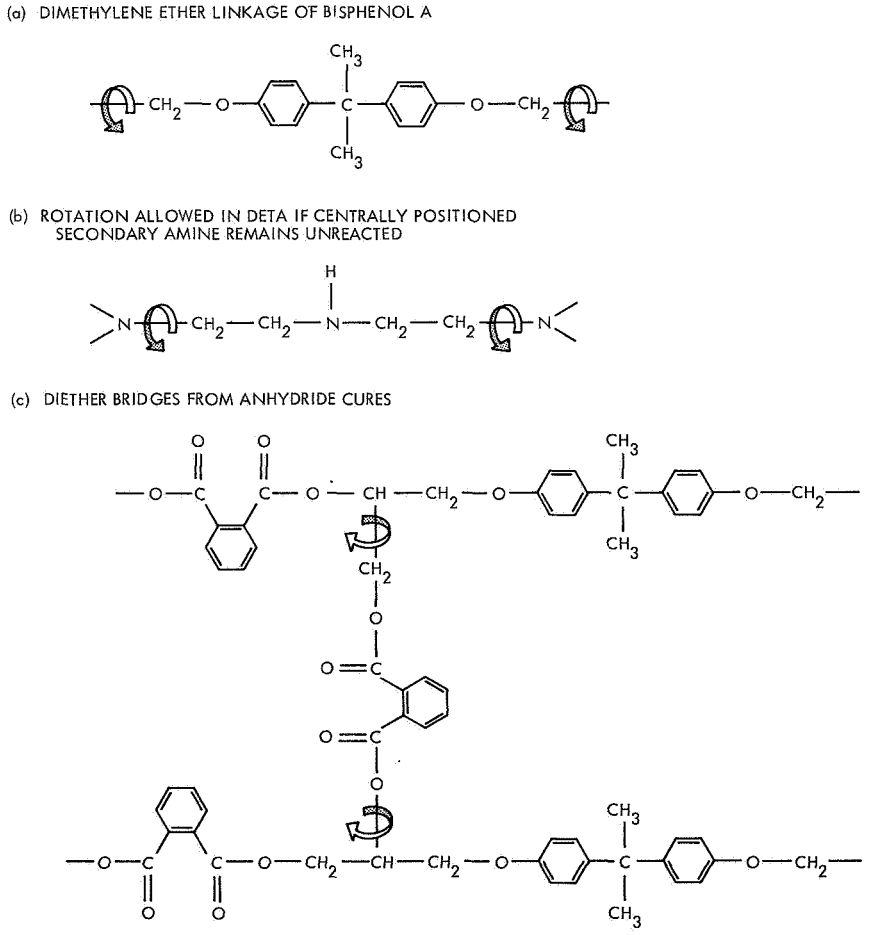


Fig. 5. Mobile groups in cured epoxy resins

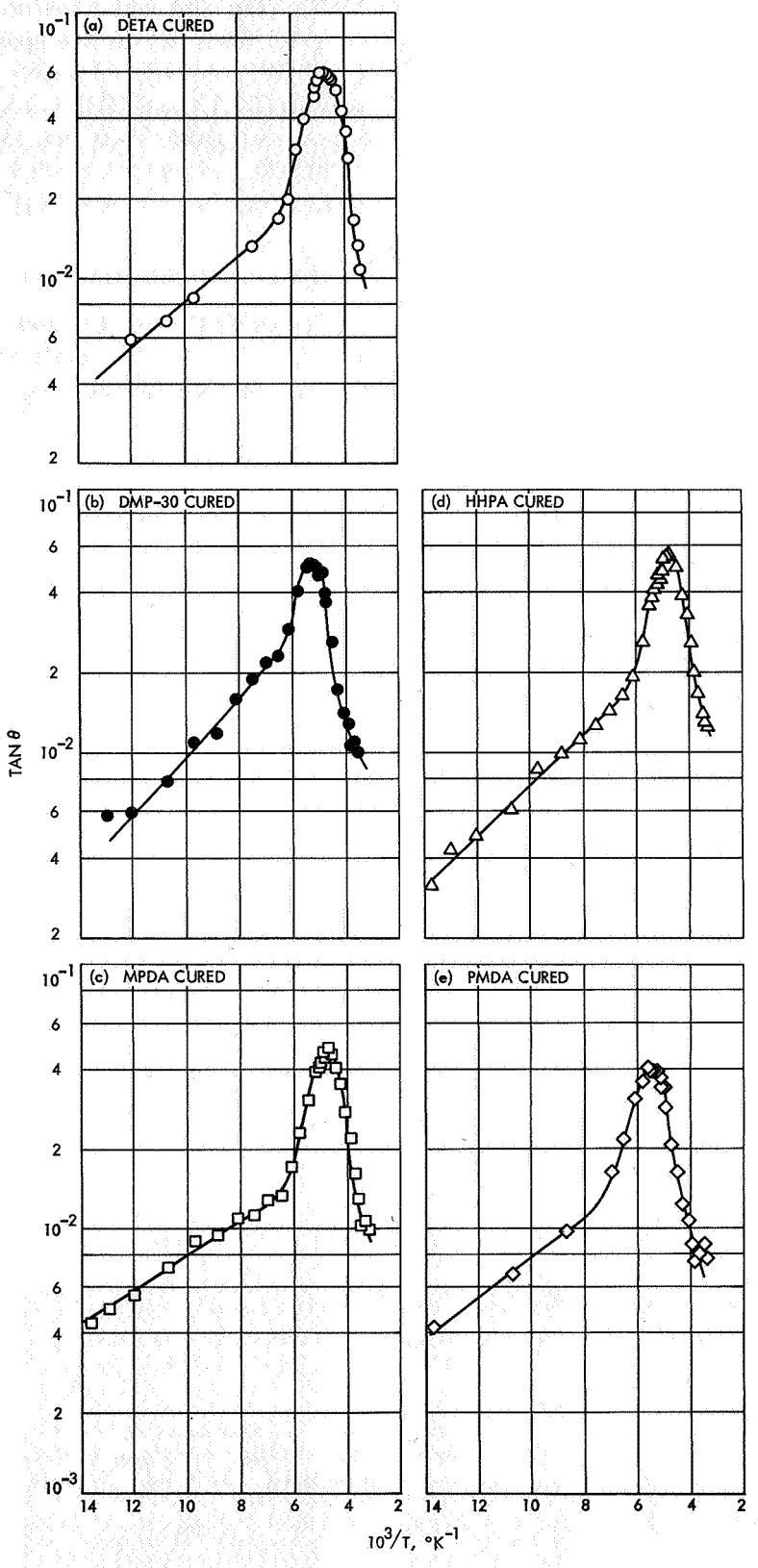


Fig. 6. Loss tangents for the cured epoxy resins

linear dependence of the logarithm of tan θ with reciprocal temperature. Aside from the β transition, no other transitions are observed in the curves, especially in the case of DETA and HHPA where an additional transition was anticipated as each had an additional mobile segment. However, the temperatures corresponding to the maximum of the β transition, as well as the values of tan θ at the maximum of the β transition, vary for the five systems (Table 2).

4. β Transition

Since DGEBA is responsible for the β transition, then the concentration of DGEBA should be related to the intensity of the β transition. Figure 7 shows plots of tan θ for each epoxy divided by its respective concentration of DGEBA. It is apparent that these DGEBA-normalized curves divide into two distinct sets, one of which consists of three epoxies in which the normalized β transitions superimpose (after shifting along the temperature axis), and the other consisting of two epoxies whose curves could not be superimposed. The two

Table 2. β transition properties of the cured epoxy resins

Curative	T_{max}, °C	$(\tan \theta)_{max}$	Curative functionality
DETA	−68	0.0626	2
DMP-30	−85	0.0525	3
MPDA	−64	0.0489	2
HHPA	−65	0.0563	2
PMDA	−95	0.0415	4

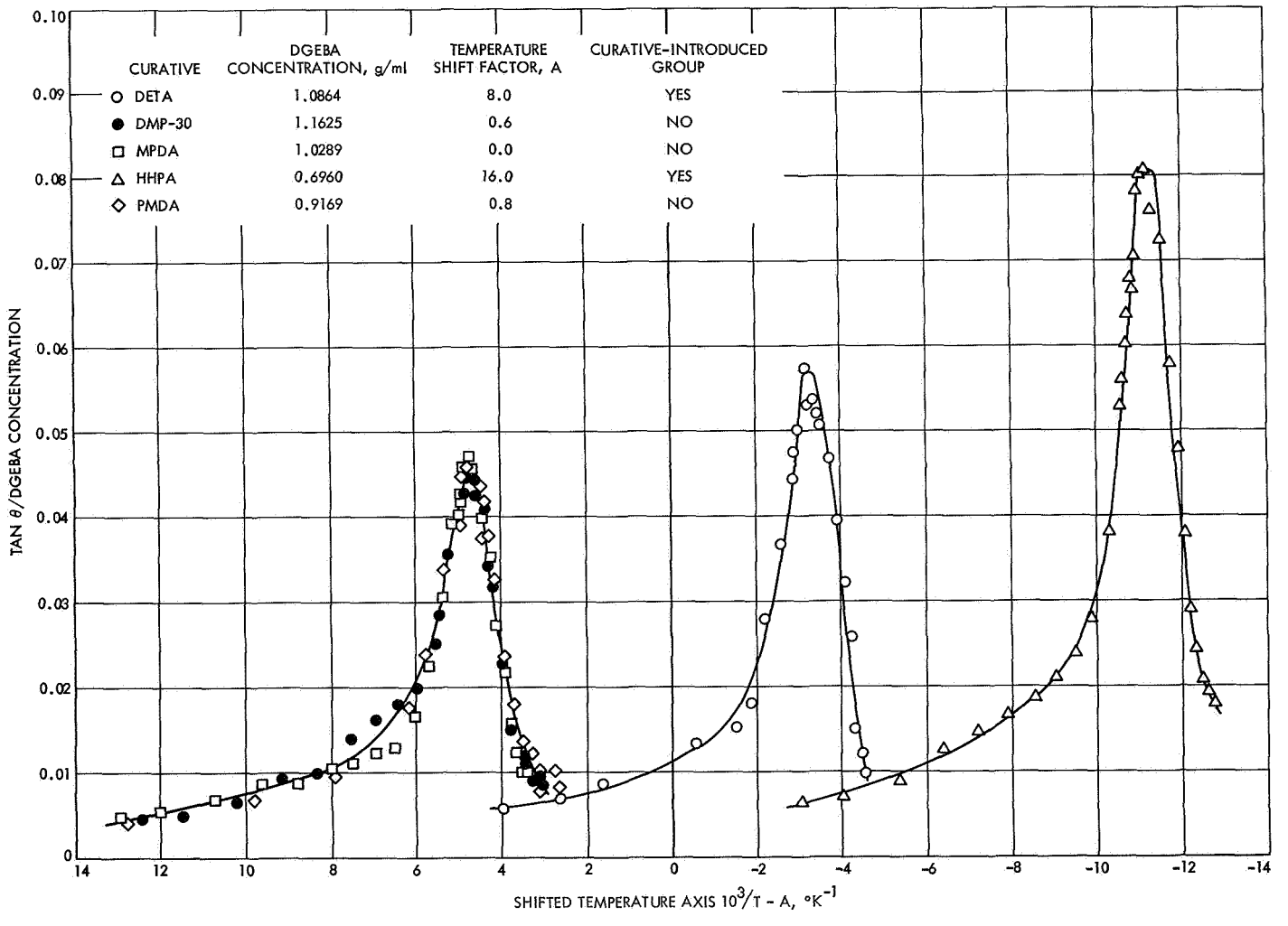


Fig. 7. Loss tangents for the cured epoxy resins normalized on the basis of DGEBA concentration

epoxies not yielding superimposed normalized β transitions were DETA and HHPA, each having a curative-introduced group, whereas the three which superimposed contained only the common DGEBA-introduced group. The intensity of the β transition for both DETA and HHPA, when normalized on a DGEBA basis, exceeded that for the other three, which strongly suggests that curative-introduced groups in epoxies having mobility act to enhance the activity of DGEBA involvement in the β transition. This is of significance since it is currently accepted that impact resistance and other shock-related properties of polymers in their solid state are related to the β transition.

5. Standard Linear Viscoelastic Model

Although the five epoxy curves did not all superimpose when normalized to the concentration of DGEBA, when plotted as the logarithm of tan θ versus reciprocal temperature, they could be shifted to a common master curve (Fig. 8). The bell shape of this master curve can be fitted with a mathematical expression derivable from the so-called standard linear viscoelastic model (Ref. 1), which has a single relaxation time. Such a restriction is generally assumed for the mobility of localized and isolated groups along the backbone of a polymer below its glass transition temperature. The expression is

$$\tan\theta = 2(\tan\theta)_{\max}\frac{\exp\left[\frac{H}{R}\left(\frac{1}{T}-\frac{1}{T_m}\right)\right]}{1+\exp\left[\frac{2H}{R}\left(\frac{1}{T}-\frac{1}{T_m}\right)\right]}$$

where

$\tan\theta$ = value of the loss tangent of the β transition

$(\tan\theta)_{\max}$ = maximum value of the β transition

H = activation energy of the β transition

T = absolute temperature in °K

T_m = temperature at the maximum

R = gas constant

The fit of this expression to the data (Fig. 8) is seen to be excellent. The activation energy for the bell-shaped master curve was found to be $H = 2.88$ kcal. Unfortunately, no corroborating activation energy data was found on other epoxies or on the other DGEBA-containing polymers such as polycarbonate and polysulfones. As all five epoxy curves superimposed, the activation energy associated with the DGEBA process

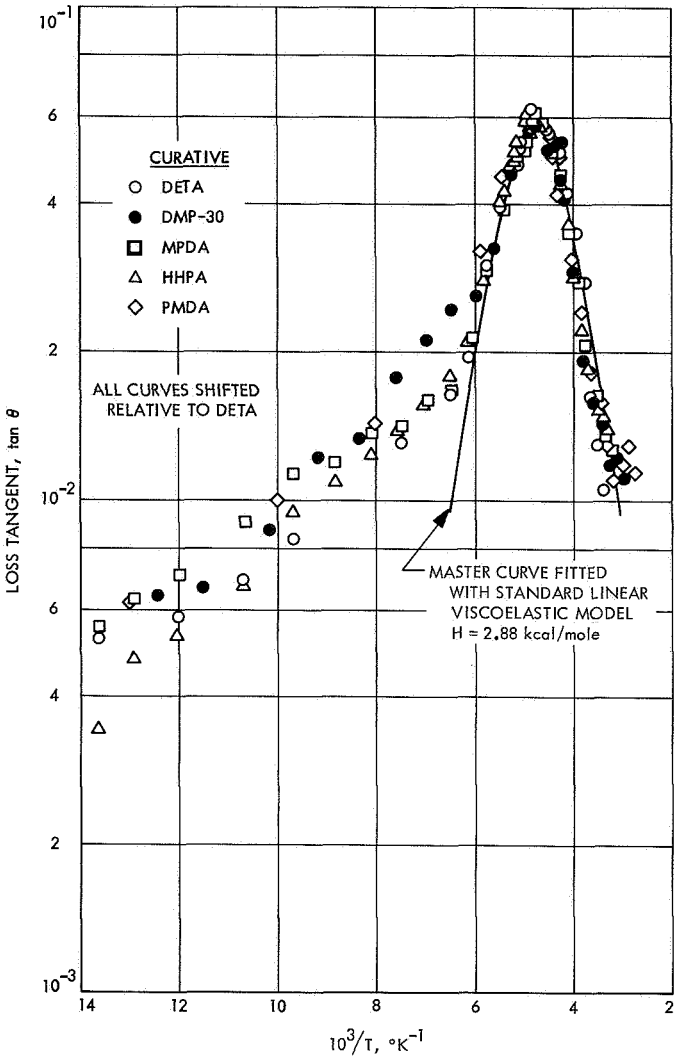


Fig. 8. Loss tangent master curve

in each of the different epoxies is the same, inasmuch as the activation energy is only related to the half-width of the bell-shaped curve and is independent of $(\tan\theta)_{\max}$ and T_m. This implies that the curative segments did not cause a new transition but only enhanced the intensity of the DGEBA process. Based on the direction of this analysis, it must be concluded that the diether bridges from anhydride curves (Fig. 5c) strongly enhance the DGEBA β transition, possibly also explaining the high impact values observed for phthalic-anhydride cured epoxies (Ref. 2).

6. Temperature of the β Transition Maximum

As previously mentioned, the temperature at the maximum of the β transition varied for the five epoxy resins from −95°C for PMDA, to −85°C for DMP-30, to

essentially a constant −65°C for DETA, MPDA, and HHPA (Table 2). No reasonable explanation for this has yet been developed; however, the correlation between the functionality of the curative and the temperature (Table 2) may provide a clue. The PMDA has a functionality of four and the lowest temperature, DMP-30, with a functionality of three, is next higher, followed by a cluster of the three remaining curatives at −65°C each having a functionality of two.

At first inspection, it appears that DETA is properly a trifunctional curative and should therefore have exhibited a temperature comparable to the other trifunctional curative DMP-30. However, if the central amine of DETA had reacted during cure, there would have been no mobile curative-introduced group for DETA. This would have resulted in the β transition of that system simply superimposing based on normalization by DGEBA concentration. The failure of the DETA system to perform in this manner, coupled with the temperature at maximum clustering with two definite difunctional curatives, provides strong dynamic-analysis evidence for the failure of the central amine to have reacted. This conclusion is supported by some indirect chemical evidence given in Ref. 3.

Returning again to the correlation between temperature at the maximum and functionality, and if we further argue that the DGEBA process requires a minimum free-volume in which to occur, then perhaps the functionality of the curative is in some way involved in establishing the level of free volume in polymers.

7. Stretch-Induced Transitions

In examining the Hirschfelder Molecular Models for the mobile segments, it was concluded that the DGEBA-introduced group had mobility only if the backbone chains were in a stretched or extended configuration, i.e., bending, coiling, or kinking of the chain at this point prevented free mobility of the DGEBA group. This was confirmed independently by O. Katz and A. V. Tobolsky (Ref. 4) from modulus measurements related to the configuration of the backbone. Accepting this point of view, it follows immediately that the intensity of the β transition might be increased during dynamic testing if the sample is under tension.

References

1. McCrum, N. G., Read, B. E., and Williams, G. *Anelastic and Dielectric Effects in Polymeric Solids*, Chapter 4. John Wiley & Sons, Inc., New York, 1967.
2. Lee, H., and Neville, K., *Epoxy Resins*, p. 121. McGraw-Hill Book Co., Inc., New York, 1957.
3. Lee, H., and Neville, K., *Handbook of Epoxy Resins*, Chapter 6, p. 7. McGraw-Hill Book Co., Inc., New York, 1967.
4. Katz, O., and Tobolsky, A. V., *Polymer*, Vol. 4, p. 417, 1963.

C. Simulation of the Venus Atmosphere, *R. H. Silver, S. H. Kalfayan, and V. W. Bellin*

1. Introduction

The Venus atmosphere was simulated for the purpose of an initial study of the effects of planetary environments on the properties of polymeric materials. The simulation was based on the information obtained by the planetary landing probe *Venera 4* (*Section D*).

2. Composition of the Simulated Atmosphere

The Venus atmosphere is reported to be composed as follows (Refs. 1 and 2):

CO_2	90 ±10% (probably > 90%)
O_2	0.4–1.6% (probably ~ 1%)
N_2	< 7% (probably < 2.5%)
H_2O	0.1–0.7%

Based on the above data, a gas mixture of the following percent by weight composition was ordered:

O_2	0.75%
H_2O	1.1%
CO_2	98.15%

The percentage of water was set at 1.1 to ensure a sufficient supply of water vapor. However, at the specified bottle pressure of 600 psig, and at ambient temperature, a gas mixture containing 1.1% water vapor was not expected.

The following mass-spectrographic analysis is typical of the composition of the gaseous mixture sampled directly from the supply bottles:

O_2	0.73%
H_2O	0.33%
CO_2	98.8%
N_2	0.11%
Ar	0.01%

Analysis of the above gas mixture, sampled after it had been in the test chamber operating at test conditions, showed the following composition:

O_2	0.84%
H_2O	0.07%
CO_2	98.6%
N_2	0.49%
Ar	0.1%

The increase in the percentage of oxygen, nitrogen, and argon, the normal constituents of air, is probably due to imperfect flushing of the test chamber by the gas mixture.

3. The Temperature and Pressure of the Simulated Atmosphere

Venera 4 indicated the Venus atmosphere to be 18.5 ±3.5 atm and the temperature to be 544°K (520°F) at the surface (Refs. 1 and 3). Although these figures are disputed on the ground that *Venera 4* probably had not reached the surface when it stopped sending data (Refs. 3 and 4), values close to those obtained by *Venera 4* were used in the present experiments; the temperature and pressure were set at 550 ±10°F and 260 ±5 psig, respectively.

4. Description of the Test Chamber

The chamber that served as the simulator was a 10-in.-diam × 12-in.-long cylindrical tank constructed of stainless steel. Ports on each side of the cylinder served for

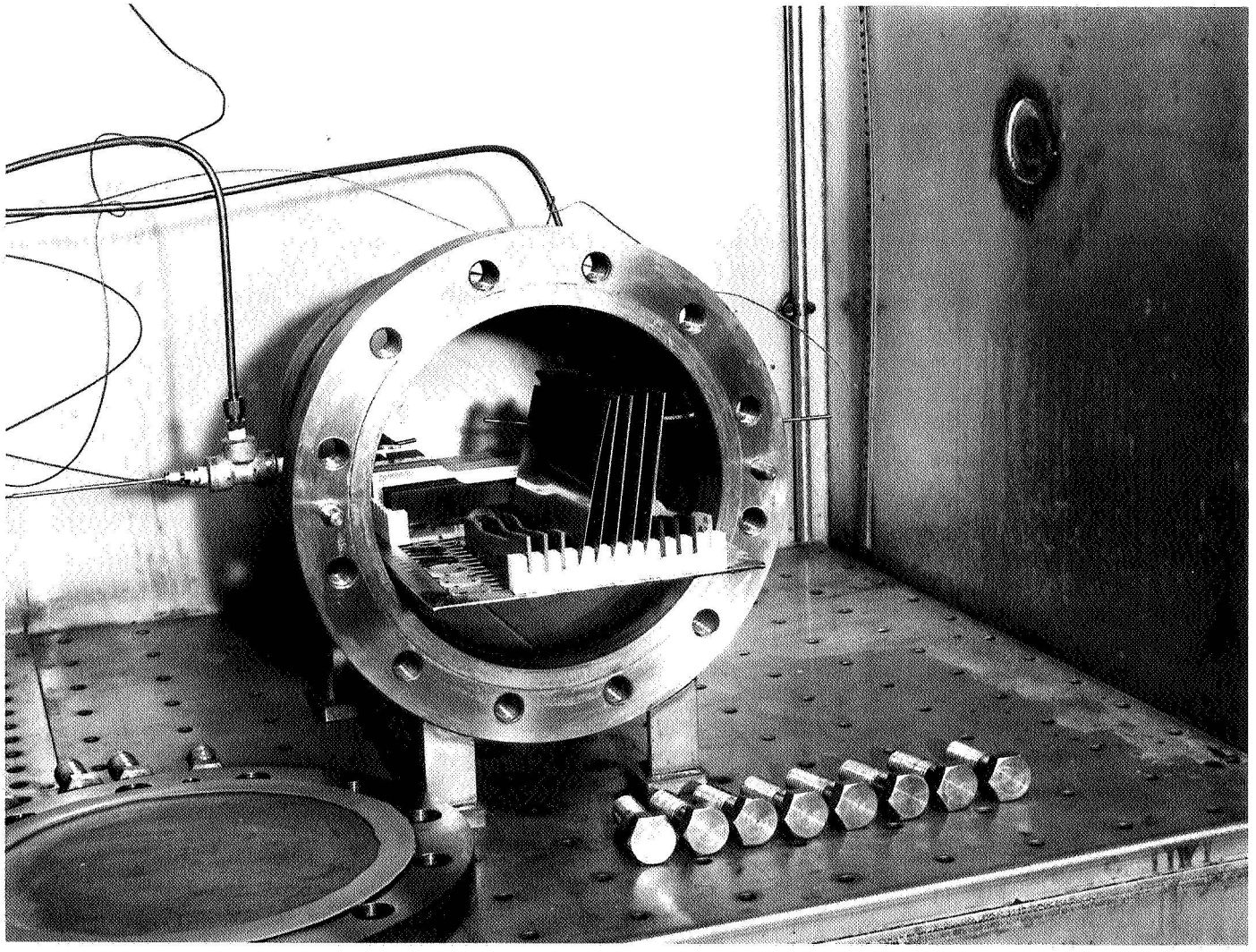


Fig. 9. Simulator test chamber

gas inlet and outlet. Stainless-steel sheathed copper–constantan thermocouples were inserted near each port. The front cover plate, as originally supplied by the manufacturer, had an asbestos–graphite gasket seal, which proved inadequate at the test pressure because of leakage. This problem was solved by using a ⅛-in.-thick teflon disc with the same diameter as the cover plate. The disc could be used for several runs at the test conditions without significant distortion or deterioration. Test specimens of polymeric materials were placed on two stainless-steel shelves within the cylindrical tank (Fig. 9).

The simulator was placed in a Conrad environmental chamber capable of heating it to the desired temperature and holding the temperature within the specified tolerance. The simulator was then evacuated three successive times and purged with the test gas to eliminate residual air. During the fourth and final purging, the gas pressure was brought to 50% of the required value. The steadily rising gas pressure, resulting from heating of the environmental chamber, was adjusted to its desired value after the test temperature of 550 ±10°F was reached. Gas flow through the simulator, at a rate of 10 ml/min, was carried on during the entire course of an exposure. Sampling for mass-spectrographic analysis was made after externally reducing the pressure in the simulator (Fig. 10). Thermocouples, the outputs of which were recorded on a potentiometer strip-chart recorder, were placed in both the simulator and environmental chamber and were used to constantly read out the temperature during a run.

Figure 11 shows the time–temperature–pressure conditions that existed during a run. A cam-type timer was set to turn off the heat automatically after the required exposure period of 6, 24, or 72 h.

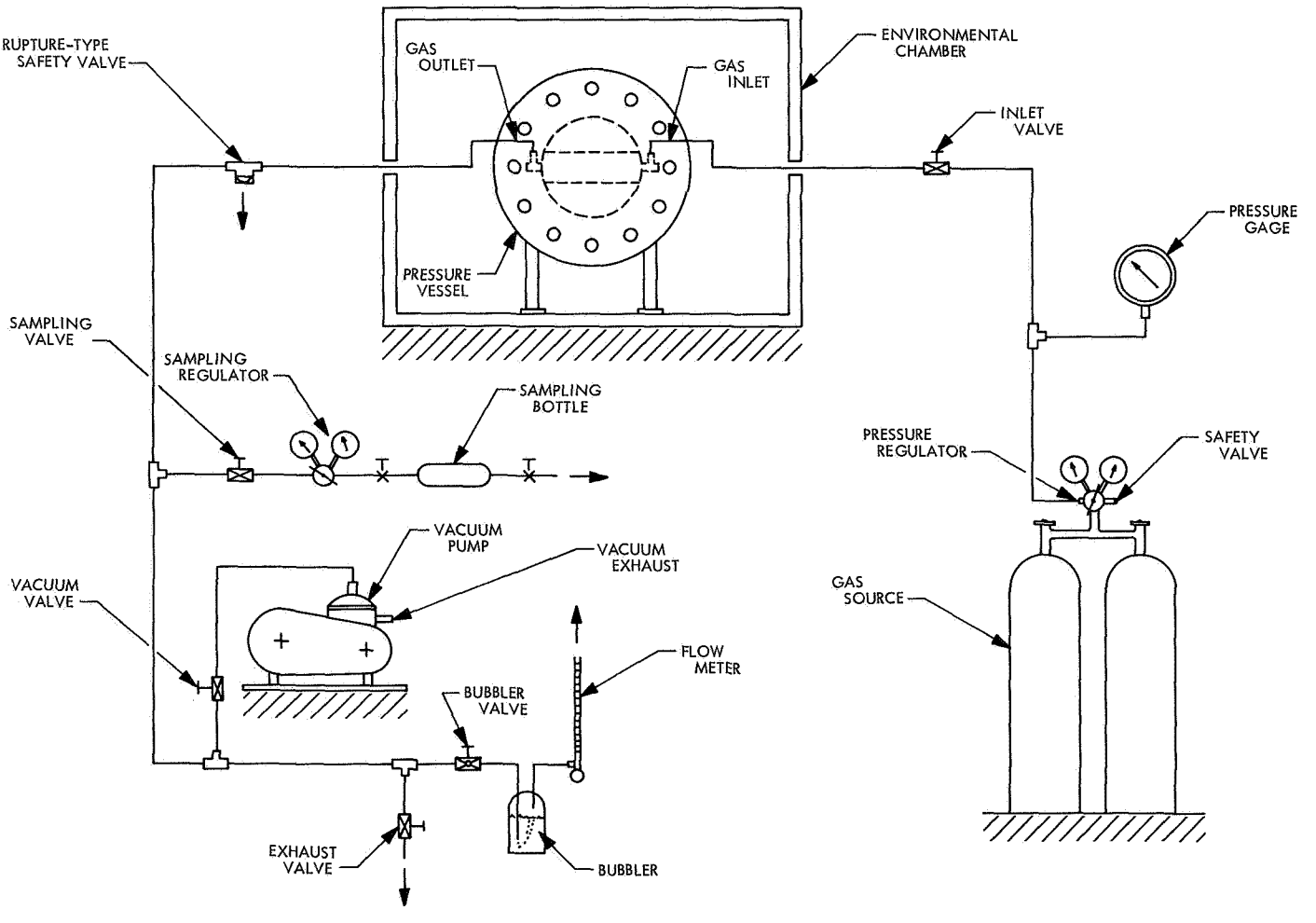


Fig. 10. Test setup diagram

References

1. Davydov, V. D., "New View of Venus," *Sci. Technol.*, pp. 49–54, Jan. 1968.
2. Vinogradov, A. P., Surkov, U. A., and Florensky, C. P., *J. Atm. Sci.*, Vol. 25, No. 4, pp. 535–536, July 1968.
3. Johnson, F. S., "The Atmosphere of Venus: A Conference Review," *J. Atm. Sci.*, Vol. 25, No. 4, pp. 658–662, July 1968.
4. Jastrow, R., "The Planet Venus," *Science*, Vol. 160, No. 3835, pp. 1403–1410, June 1968.

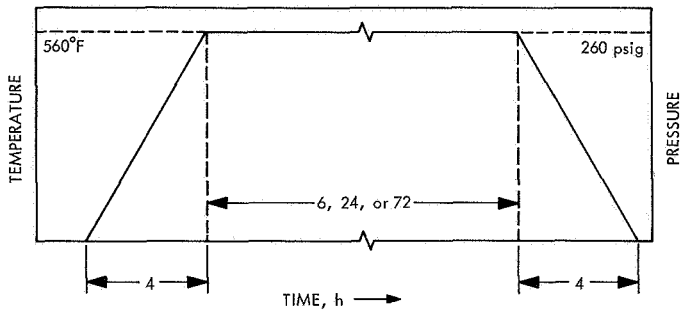


Fig. 11. Conditions during a cycle

D. Effects of Simulated Venus Atmosphere on Polymeric Materials, *S. H. Kalfayan, R. H. Silver, and V. W. Bellin*

1. Introduction

In October 1967, *Venera 4* entered Venus' atmosphere and sent back data concerning temperature, pressure and atmospheric composition as it descended to the surface (Ref. 1, 2). Based on this data, environmental test conditions were set up to simulate the Venus atmosphere (*Section C*) and a number of selected polymeric materials were tested after exposure to this environment. All tested materials were commercially available products with proven compatibility to spacecraft sterilization conditions. They also represented a number of functional classes, such as adhesives, encapsulants, coatings, circuit board materials, films, etc.

This article summarizes the results obtained from these tests.

2. Experimental

a. Environmental test conditions. The test equipment used is described in *Section C*. Table 3 shows the test conditions used in the Venus atmospheric simulation. These conditions, first thought to exist on the equatorial surface of Venus, most probably describe the conditions at some distance from the surface. There are arguments that support the hypothesis that *Venera 4* did not reach the planet's surface before it stopped sending data (Ref. 3).

b. Sample materials and preparation for testing. As previously mentioned, the materials tested were commercial products of well-known polymeric composition. Test specimens were prepared as required by the standard test methods used, i,e., some of the products (e.g., the encapsulants and adhesives) required such preliminary handling as mixing and degassing before castings or test specimens could be prepared.

c. Tests used. Standard tests were used as applicable to each class of compounds. Weight losses were measured to an accuracy of ±0.1 mg and dimensional change measurements were accurate to ±0.1 mil.

3. Results and Discussion

The polymeric products of interest were subjected to a screening program consisting of only 6 h of exposure to the conditions shown in Table 3 and a limited amount of testing after exposure. The materials that passed the preliminary screening tests were subjected to the same conditions for longer durations and then tested more extensively.

The results of the preliminary screening tests are given in Table 4. This table also shows the basic polymeric constituent of each product and the functional class to which it belongs. From this table, it may be seen that only six products met the compatibility criteria set for the screening program.

Two materials, Teflon FEP (polyperfluoroethylene propylene, No. 14. Table 4) and Exp 820 (poly-2,2′-hexamethylene-5,5′-bibenzimidazole, No. 15, Table 4) had melted and deformed at the exposure temperature. However, after the specimens were brought to room conditions, the mechanical properties seemed intact and the little weight loss indicated that chemical degradation

Table 3. Environmental test conditions

Temperature	550 ±10°F
Pressure	18.5 ±1 atm
Gaseous atmosphere	
Carbon dioxide	96.7–99.5%
Water vapor	0.1–0.7%
Oxygen	0.4–1.6%
Nitrogen	<1%

was practically absent. If these two plastics had been reinforced by fibrous materials, they probably would have been dimensionally stable. In the near future, to test this hypothesis, a glass-reinforced and a graphite-reinforced polybenzimidazole will be exposed to the same conditions.

The six products that passed the preliminary screening were then exposed for 24 h to the simulated Venus atmosphere. Three of these (i.e., Viton 77-545, GP-77 and Pyre-ML, RK 692) failed to meet the compatibility criteria, which essentially consisted of retaining at least 70% of original mechanical, physical, and electrical properties and losing less than 2% of original weight (Tables 5 and 6). The three remaining compounds (Teflon TFE, Fiberglas 91LD, and Kapton 100) then were exposed 72 h to the simulated conditions. Kapton 100, the sample most affected by this exposure, retained 40% of its original elongation and 63% of its original tear strength after this treatment. Its tensile strength was well retained. The other two compounds were affected very little by the 72-h exposure.

Table 4. Results of preliminary screening tests[a]

No.	Commercial designation	Material type	Functional class	Comments
1	DEN 438/amine	Epoxy Novolac/amine	Adhesive	Test samples came apart
2	RTV 891	Silicone	Adhesive	Very low tensile shear strength
3	RTV 140	Silicone	Adhesive	Test samples came apart
4	GP-77	Silicone resin	Coating	Pass
5	Pyre-ML, RK 692	Polyimide	Coating	Pass
6	Viton 77-545	Fluorocarbon	Elastomer	Pass
7	Hadbar 5000/50	Fluorosilicone	Elastomer	Melts
8	PMP 6035	Silicone	Elastomer	Embrittlement
9	Sylgard 184	Silicone	Encapsulant	Softens
10	Kapton 100	Polyimide	Film	Pass
11	EG 758	Epoxy/glass	Reinforced plastic	Charring
12	Fiberglas 91LD	Phenolic/glass	Reinforced plastic	Pass
13	Teflon TFE	Fluorocarbon	Plastic	Pass
14	Teflon FEP	Fluorocarbon	Plastic	Melts
15	Unfilled Exp 820	Polybenzimidazole	Plastic	Softens and deforms
16	PPO	Polyphenzlene oxide	Filled plastic	Decomposes

[a]Tests performed after 6-h exposure to the simulated Venus atmosphere.

Table 5. Changes in mechanical properties after exposure to a simulated Venus atmosphere

Material	Tensile strength retained, %			Elongation retained, %			Hardness retained, %		
	After 6 h	After 24 h	After 72 h	After 6 h	After 24 h	After 72 h	After 6 h	After 24 h	After 72 h
Teflon TFE	105	106	109	102	104	104	105	101	100
Fiberglas 91LD	94	96	91	97	94	94	106	108	103
Kapton 100	92	91	89	91	54	40		89[a]	63[a]
Viton 77-545	92	50	—	91	40	—	93	105	—
GP 77	—	—	—	—	—	—	80[b]	Peeling started	Not tested
Pyre-ML, RK 692	—	—	—	—	—	—	70[b]	>20[b]	

[a]Property measured was tear strength.
[b]Property measured was scrape-adhesion.

Table 6. Changes in physical and electrical properties after exposure to a simulated Venus atmosphere

Material	Weight change, %			Dimensional change, %			Volume resistivity, Ω-cm			Dielectric strength, V/mil		
	After 6 h	After 24 h	After 72 h	After 6 h	After 24 h	After 72 h	Control	After 24 h	After 72 h	Control	After 24 h	After 72 h
Teflon TFE	+0.11	+0.19	+0.30	+0.80	+0.814	+1.533	6.3×10^{15}	6.0×10^{15}	1.46×10^{16}	339	320	335
Fiberglas 91 LD	−1.40	−0.148	−1.88	−0.290	none	−0.200	1.44×10^{14}	1.03×10^{14}	7.3×10^{14}	411	437	378
Kapton 100	+0.99	+1.95	+0.780	—	—	—	1.24×10^{17}	1.20×10^{17}	1.16×10^{17}	3740	3745	3837
Viton 77-545	−0.67	−2.64	—	−1.91	−3.80	—	—	—	—	—	—	—
GP 77	—	Not tested	—	—	Not tested	—	—	—	—	—	—	—
Pyre-ML, RK 692	—	Not tested	—	—	Not tested	—	—	—	—	—	—	—

4. Conclusion

Only a few of the polymeric materials presently used in spacecraft would be able to withstand a simulated Venus atmosphere, and, the conditions at the surface of Venus are probably more severe than those used in the simulation.

References

1. Davydov, V. D., "New View of Venus," *Sci. Technol.*, pp. 49–54, Jan. 1968.
2. Vinogradov, A. P., Surkov, U. A., and Florensky, C. P. "The Chemical Composition of the Venus Atmosphere," *J. Atm. Sci.*, Vol. 25, pp. 535–536, July 1968.
3. Johnson, F. S., "The Atmosphere of Venus: A Conference Review," *J. Atm. Sci.*, Vol. 25, pp. 658–662, July 1968.

E. Poly(isobutylene) Prepolymers, *J. Miller*

1. Introduction

Solid propellant prepolymers must exhibit minimum reactivity with propellant oxidants, be of maximum average molecular weight and low bulk viscosity, and yield cured solid propellants with relatively temperature-independent mechanical properties (SPS 37-45, Vol. IV, pp. 113–115). Difunctional poly(isobutylene) (PIB) seems to satisfy the above requirements; however, its limited means of preparation has discouraged the use of this polymer as a solid-propellant prepolymer candidate (SPS 37-45, Vol. IV). Isobutylene is readily polymerized by strong Lewis acids A in the presence of minute quantities of H_2O to yield a terminally-unsaturated, monofunctional polymer with a degree of polymerization $\overline{X}_n$ inversely proportional to the temperature of polymerization. The mechanism of polymerization is believed to be

$$A + H_2O \rightleftharpoons AOH^- H^+$$

$$AOH^- H^+ + CH_2 = C(CH_3)_2 \quad (M) \xrightarrow{k_i} HM^+ AOH^-$$

$$HM^+ AOH^- + nM \xrightarrow{k_p} CH_3-C(CH_3)_2-(CH_2-C(CH_3)_2)_n^+ \; AOH^-$$

$$CH_3-C(CH_3)_2-(CH_2-C(CH_3)_2)_n^+ \; AOH^- \xrightarrow{k_t} CH_3-C(CH_3)_2-(CH_2-C(CH_3)_2)_{n-1}-CH_2-C(CH_3)_2 \quad (P) + AOH_2$$

$$CH_3-C(CH_3)_2-(CH_2-C(CH_3)_2)_n^+ \; AOH^- + M \xrightarrow{k_{tr}} P + HM^+ AOH^-$$

where

k_i = rate constant for initiation of polymerization

k_p = rate constant for propagation of polymerization

k_t = rate constant for termination of polymerization

n = number of monomer molecules or units

M = monomer

P = polymer

k_{tr} = rate constant for chain transfer

When rate termination R_t predominates, the following relationship is valid:

$$\bar{X}_n = \left[\frac{R_p}{R_t}\right] M$$

where R_p is the rate of polymerization. Thus, if the activation energy for termination is greater than that for propagation, the polymer molecular weight will increase with decreasing temperature. It can be seen from the above that if an unsaturated Lewis base is substituted for H_2O, and if $R_t \gg R_{tr}$ (rate of transfer), it might be possible to prepare a difunctional PIB with a number average molecular weight $\bar{M}_n$ dependent on temperature and monomer concentration. A program was thus initiated to prepare PIB using $TiCl_4$, a strong A, and allylic halides as unsaturated Lewis bases.

2. Experimental

a. Equipment. The desired mechanism for initiation using $TiCl_4$ and allylic halides (e.g., allyl chloride) is

$$TiCl_4 + CH_2 = CH - CH_2Cl \longrightarrow CH_2 = CH - CH_2^+ \, TiCl_5^-$$

$$CH_2 = CH - CH_2^+ \, TiCl_5^- + CH_2 = C(CH_3)_2 \xrightarrow{R_i} CH_2 = CH - CH_2 - CH_2 - C(CH_3)_2 + TiCl_5^- \xrightarrow{\text{etc.}}$$

Any water or air present in the system would compete with allyl chloride and decrease the overall functionality of the polymer. Such competition necessitated the construction of a high-vacuum system and an efficient procedure to dry all reagents (Fig. 12). West-glass high-vacuum, greaseless stopcocks were used to minimize solution of isobutylene, allylic halides, and hydrocarbon solvents in the conventional stopcock greases. A vacuum of 1×10^{-8} torr could be attained in 72 h by operating the ion and diffusion pumps in series. The vacuum could then be maintained, and occasionally improved, with the Ultek ion pump. This procedure eliminated baking of the system and constant monitoring of the traps on the diffusion pump. An accurate reading of the system's pressure could also be obtained from the Ultek gage thus eliminating the use of troublesome ionization gages.

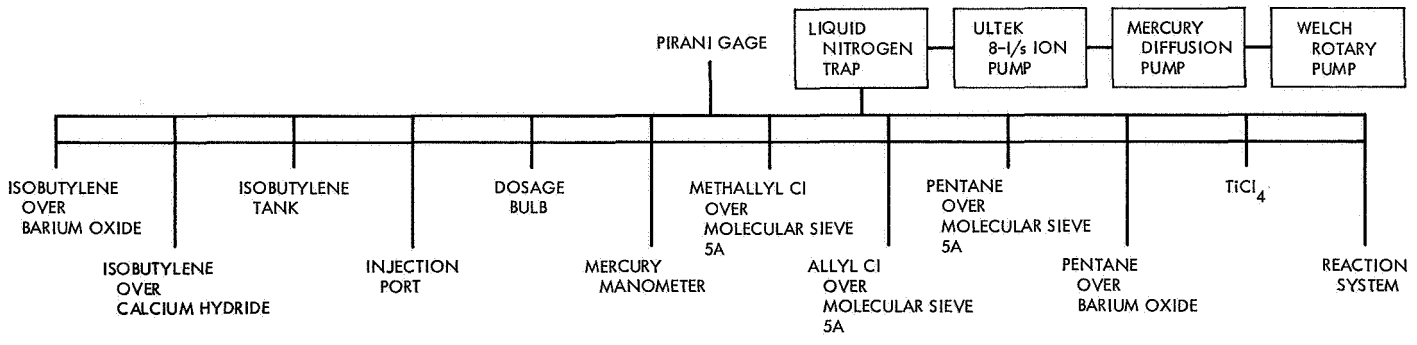


Fig. 12. Vacuum system for isobutylene polymerization

b. Purification of reactants. The $TiCl_4$ (boiling point = 136°C) was distilled at atmospheric pressure under nitrogen. The middle fraction was collected and introduced into the vacuum system at the injection port and collected at dry ice–acetone temperature. The $TiCl_4$ was distilled twice more on the vacuum train at ambient temperature and collected at dry ice–acetone temperature. The middle fraction was kept after each distillation.

Pentane was distilled at atmospheric pressure under nitrogen onto a 5A Linde molecular sieves. The middle fraction was kept and introduced into the vacuum system, collected at liquid-nitrogen temperature over a molecular sieve, stirred and degassed, and finally distilled onto anhydrous barium oxide.

Allyl chloride and methylallyl chloride was purified analogously to pentane and stored over a 5A molecular sieve.

Isobutylene was introduced at its injection port onto calcium hydride at liquid-nitrogen temperature, warmed to room temperature, stirred for a week, and finally distilled onto and stored over barium oxide.

All reagents were periodically stirred over their drying agents for several days and degassed to 10^{-6} torr at − 196°C.

c. Procedures and techniques. Polymerizations were run in a specially-constructed reaction vessel (Fig. 13). All reagents were degassed to 10^{-6} torr, followed by evacuation of the entire system to a pressure of 10^{-8} torr. A 5-ml sample of methanol was sealed in the section indicated in Fig. 13. Standard vacuum-line procedures were used to manipulate the reactants (Ref. 1). The required amounts of isobutylene and allyl chloride were measured into the dosage bulb and condensed in section B. The desired amount of pentane was measured into section A and then condensed into section B. Finally, the $TiCl_4$ was condensed and measured into section A followed by sufficient pentane to increase the total volume to 10 ml. The reaction system was sealed, removed from the vacuum line, placed in a Conrad chamber and allowed to equilibrate at the desired temperature for 3 h. The break seal on section A was broken and the vessel was inverted allowing the $TiCl_4$ solution to flow into section A. The resultant yellow solution was stirred the desired period of time. Polymerization was finally terminated by destruction of the catalyst with methanol. The solution was washed with water and dried

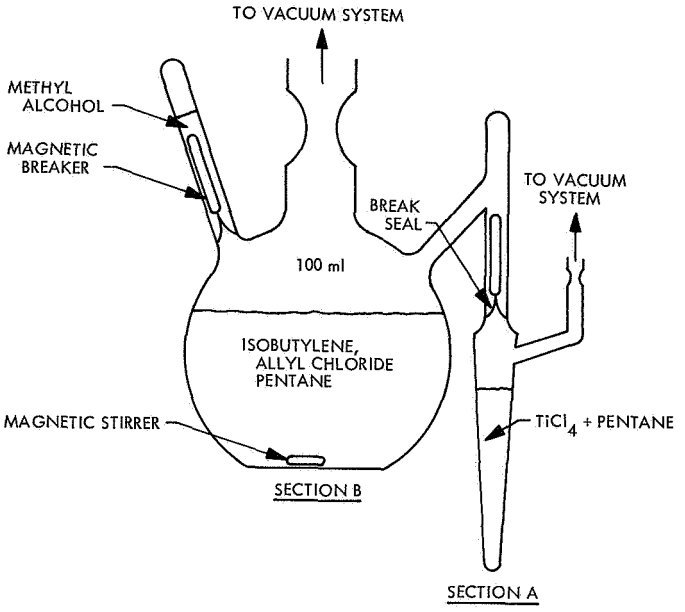


Fig. 13. Polymerization reaction vessel

over magnesium sulfate. The pentane was removed at reduced pressure and the resultant polymer was dried at 50°C under vacuum.

Intrinsic viscosities were determined at 25°C in pentane by standard procedures. Viscosity average molecular weights were determined from the equation

$$[n] = K'\overline{M}_v\,\alpha$$

where $K' = 2.6110^{-4}$ and $\alpha = 0.70$ at 25°C. Number average molecular weight $\overline{M}_n$ was then determined using the relationship

$$\overline{M}_n = \frac{\overline{M}_v}{1.83}$$

Infrared spectra were recorded on a Perkin Elmer 421 spectrometer.

3. Results and Discussion

Results of nine runs will allyl chloride and one run with methallyl chloride are summarized in Table 7. The lowest molecular weight obtained was 7300 (run 6). Attempts to decrease molecular weight by increasing reaction temperature failed (run 7); however, it appears $\overline{M}_n$ is proportional to monomer concentration (compare runs 5 and 9). The single run with methallyl chloride at 0°C produced a polymer with a relatively high molecular weight.

Table 7. Polymerization test results

Run[a]	Reaction time, h	Temperature, °C	Isobutylene, mole	Allyl chloride, mole	Methallyl chloride, mole	$TiCl_4$, mole	[n]	$\bar{M}_n$
1	1.5	−30	2.58	0.0029	—	0.15	—	Very high
2	0.75	−30	2.66	0.043	—	0.13	—	Very high
3	4	−30	2.54	0.17	—	0.26	—	Very high
4	24	0	2.54	0.17	—	0.26	0.204[b]	10,360
5	24	0	2.58	0.36	—	0.52	0.503	27,400
6	48	25	2.63	0.17	—	0.27	0.200	7,300
7	96	44	2.60	0.17	—	0.27	—	No polymer
8	120	25	1.80	0.20	—	0.27	—	No polymer
9	72	0	1.36	0.34	—	0.49	0.335	15,300
10[b]	24	0	2.56	—	0.19	0.26	0.505	27,600

[a]All runs in pentane.

[b]Efficiency was higher than all other polymerizations.

Infrared spectra of samples pressed between salt plates indicated the presence of functionality; however, samples of comparable molecular weight of known functionality were not available for comparison. A nuclear magnetic resonance method is now being developed to determine functionality.

More must be done before conclusions can be drawn as to the feasibility of this method for preparing difunctional PIB prepolymers. Perhaps, through variation of temperature, time, and monomer concentration, a polymer of $\bar{M}_n$ approximating 1000 can be made. Other functional halides and more active A might be used to increase polymerization efficiency. Finally, other sources of functional initiation must not be discounted, i.e.,

$$NO_2^+BF_4,\ CH_3C(=O)^+BF_4^-,\ Et_3O^+BF_4^-,\ \text{etc.}$$

Such salts in highly polar aprotic media could yield surprising results. However, the most promising approach for obtaining difunction PIB involves direct polymerization by molecular sieves. Current work shows that difunctionally unsaturated PIB can be prepared readily by this method.

Reference

1. Sanderson, R. T., *Vacuum Manipulation of Volatile Compounds*, John Wiley & Sons, Inc., New York, 1948.

F. The Aging of Styrene–Butadiene Rubber: A Spectral Investigation, *R. Rakutis and S. H. Kalfayan*

1. Introduction

Absorption changes in the infrared spectra of polymers as the result of oxidation may be expected in the following three regions: (1) 3730–3000 cm^{-1} due to hydroxyl (−OH) groups, (2) 1840–1630 cm^{-1}, the absorption region of carbonyl (>C = O) compounds, and (3) 1350–741 cm^{-1}, which corresponds to absorption by ether, peroxide, and hydroperoxide structures (Refs. 1, 2, 3). Since a study of these changes in absorption during oxidation of styrene–butadiene rubber (SBR) appeared to provide one possible method of determining the kinetics of oxidation, it was decided to carry out a few preliminary experiments in this area. Accordingly, an exploratory investigation was made on the structural changes resulting from the oxidation of SBR films cast on sodium chloride plates. The temperature and length of time needed to effect a change in the infrared spectrum of the SBR films was observed and recorded.

It was also of interest to examine the possibility of using infrared spectroscopy as an added tool in the investigation of the changes, both mechanical and chemical, taking place in the elastomer as it was placed in a stress relaxometer at elevated temperatures.

The near-infrared region of the spectrum (1–3 μm) may also be used for the determination of certain functional groups in organic compounds. Bands at about 1.62 and 2.10 μm indicate terminal unsaturation and a band

at 2.14, cis unsaturation (Ref. 4). The primary absorption bands of the $-OH$ stretch of alcohols, hydroperoxides, and acids are found between 2.75 and 2.83 μm (Refs. 5, 6). Since there was a good possibility that the near-infrared spectra of the oxidized SBR rubber would show definite changes in the terminal methylene-group region as well as in the $-OH$ region, it was decided to look at the near-infrared spectra of the SBR films. The near-infrared also has the additional advantage that much thicker films (Ref. 7) can be used in the spectral measurements than in the case of the infrared.

2. Experimental

For the oxidative study of SBR (obtained from the National Bureau of Standards, lot 386F, type 1500), films from a 2% solution of the rubber were cast on sodium chloride plates. One sample was as-received SBR, and the other (sample 1) was milled, but not cured, with 0.9% dicumyl peroxide. After the solvent was allowed to evaporate, the sodium chloride plates were placed in an oven and heated in air at the desired temperature. The infrared spectra were taken on a Perkin-Elmer 421 recording spectrophotometer while a Cary 14 recording spectrophotometer was used for near-infrared measurements.

In preparing films for stress-relaxation measurements, SBR rubber (sample 2) was milled with 0.3% dicumyl peroxide and cured in a press for 1 h at 310°F. The desired thickness was obtained by using different sized molds in the press. (For these preliminary investigations, 7–8-mil thick films were made.) A 3 1/16-in. long dumbbell-shaped die was used to cut samples for the relaxometer. The rubber films were heated in air at 100 and 125°C, respectively, for 22 h at an extension of 16% (based on the total length of the sample).

3. Results

There were no observable changes in the infrared spectra after sample 1 was heated for 1 h at 100°C. The temperature was then increased to 120°C and the sample heated for an additional 4 h, which resulted in a gradual increase in the $>C=O$ peak (1720 cm^{-1}) and in the $-OH$ peak (3400 cm^{-1}) (Fig. 14). The ether region (1200–1000 cm^{-1}) of the spectrum also showed a general increase in the intensity of absorption. These results are in good agreement with other oxidation studies of natural as well as synthetic rubber (Refs. 8, 9, 10). The presence of the $>C=O$ and $-OH$ peaks in the spectra was explained by the decomposition of the unstable peroxides

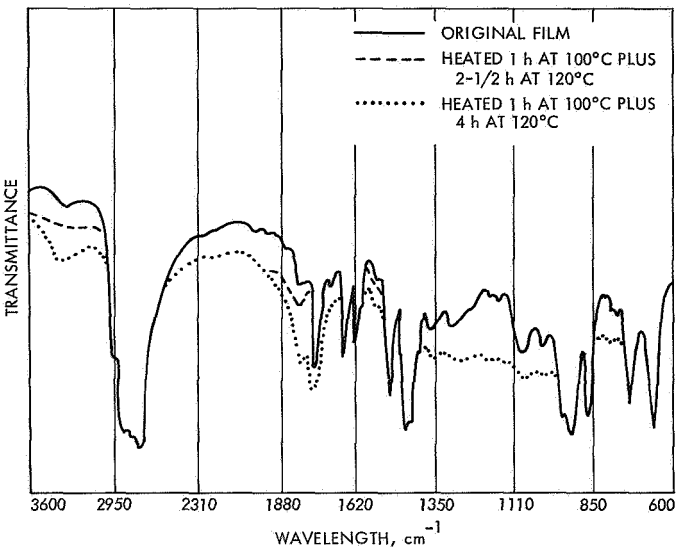


Fig. 14. Infrared spectra of SBR rubber

into a whole series of secondary reaction products, i.e., alcohols and enolized ketones (Ref. 10).

In the stress relaxation versus infrared measurements, comparisons were made of the spectrum before and after the stress relaxation experiments. Even though the 7–8-mil samples were too thick to obtain good spectra, there was enough transmission in the regions of interest to observe the changes taking place. After stress relaxation at 100°C for 22 h, the SBR film showed no spectral change other than a very slight weakening in the intensity of the small $-OH$ peak that had been present at the start. The sample had darkened in color, but was not brittle. When a new SBR sample was subjected to stress relaxation at 125°C, the film changed to a brittle, yellow material that broke towards the end of the experiment. The infrared spectrum of the exposed sample showed no detailed structural peaks. There were only three very broad regions of absorption: (1) $-OH$ at 3500–3400 cm^{-1}, (2) $>C=O$ at 1800–1700 cm^{-1}, and (3) a wide area of absorption from 1460–800 cm^{-1}.

The near-infrared spectra also showed no change after the 100°C treatment. Moreover, even after treatment at 125°C, the near-infrared spectrum of the sample was almost identical to that of the untreated spectrum (Fig. 15). There was a decrease in the areas of the two doublets at 2.36 and 2.32 μm ($-CH_2-$ absorption), but the general appearance of the spectrum was not changed. Results similar to the above were reported on natural rubber (Ref. 11). In that investigation, the chemical changes that could be measured spectroscopically by the

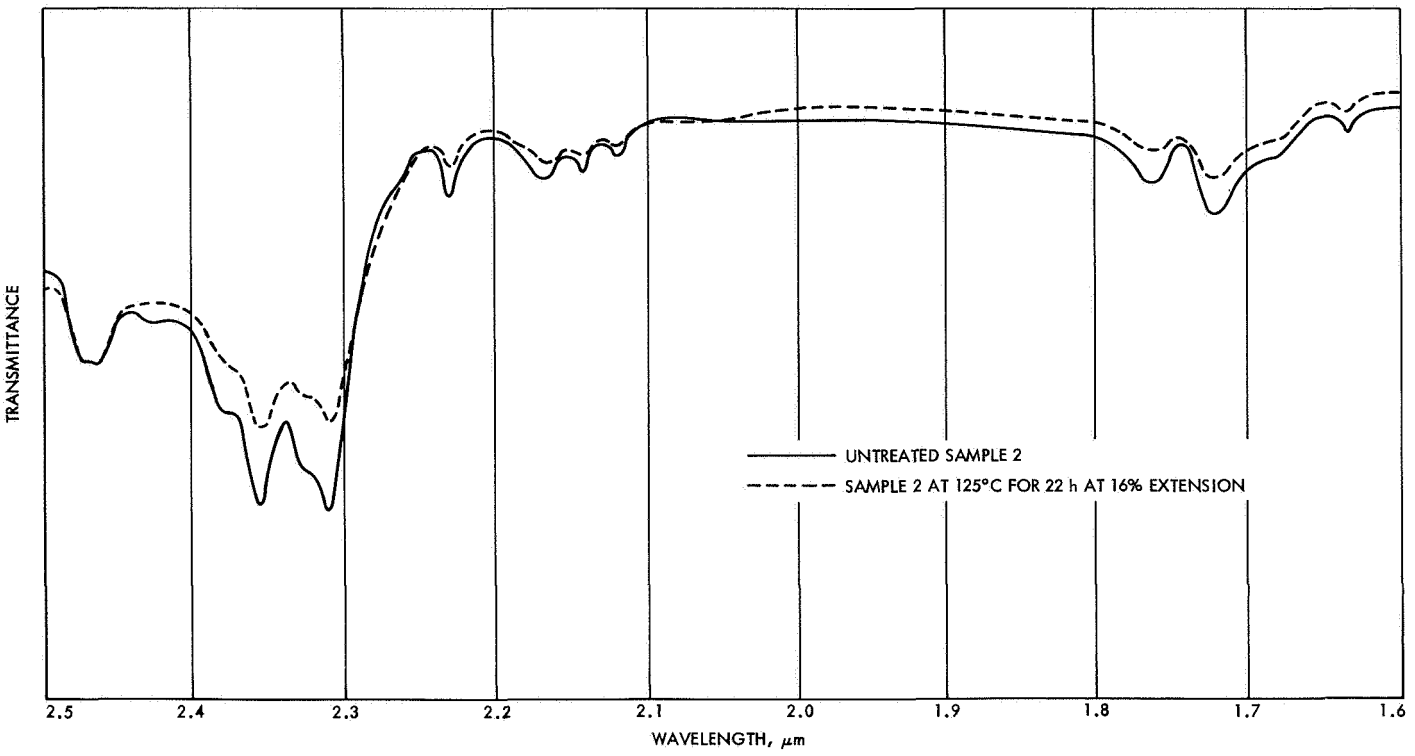


Fig. 15. Near-infrared spectra of SBR rubber

infrared occurred only at and beyond the point where the mechanical properties had already been profoundly altered. With the SBR film, another reason for the lack of change in the near-infrared spectrum was probably the fact that the $-OH$ and $\geq C = O$ stretching vibration bands lie in the 2.7 to 3.0-μm region. This region is not accessible to the Cary 14 spectrophotometer and, thus, the main products formed on autoxidation of SBR could not be observed in the near-infrared spectra.

4. Future Work

From observing the changes in the spectrum of the heated SBR film, it seems that a kinetic study could be made on the oxidation of SBR by following the change in absorbance of, for example, the $\geq C = O$ peak. However, such a study should utilize only purified SBR without anti-oxidants or other additives. The same study might be possible using the near-infrared, but only with a spectrophotometer that can be used in the 2.5 to 3.0-μm region.

Since there seems to be no clear explanation for the large concentration of $-OH$ groups formed as oxidation products, it would be interesting to attempt product isolation and subsequent identification. Perhaps more rigorous oxidation conditions would provide simpler reaction paths and a smaller number of products that could then be readily identified.

References

1. Binder, J. L., "Infrared Spectroscopy," *Rub. Chem. Technol.*, Vol. 35, p. 57, 1961.

2. Stolyarov, B. V., "Application of Infrared Spectroscopy to Studies of the Oxidation of High Polymers," *Zh. Prikl. Khim.*, Vol. 34, No. 12, p. 2726, 1961.

3. Bellamy, L. J., *The Infra-red Spectra of Complex Molecules*, John Wiley & Sons, Inc., New York, 1958.

4. Goddu, R. F., "Determination of Unsaturation by Near-Infrared Spectrophotometry," *Anal. Chem.*, Vol. 29, p. 1790, 1957.

5. Slover, H. T., and Dugan, L. R., "Application of Near Infrared Spectrophotometry to the Study of the Autoxidation Products of Fats," *J. Am. Oil Chem. Soc.*, Vol. 35, p. 350, 1958.

6. Goddu, R. F., "Determination of Phenolic Hydroxyl by Near-Infrared Spectrophotometry," *Anal. Chem.*, Vol. 30, p. 2010, 1958.

7. Willis, H. A., and Miller, R. G. J., "Quantitative Analysis in the 2-μ Region Applied to Synthetic Polymers," *J. Appl. Chem.* (London), Vol. 6, p. 385, 1956.

8. Kellö, V., Tkac, A., and Hrivikova, J., "A Study of the Aging of Natural Rubber by Means of Infrared Spectroscopy. Part V: Aging Accelerated by Heat," *Rub. Chem. Technol.*, Vol. 29, p. 1245, 1956.

9. Kellö, V., and Tkac, A., "A Study of the Aging of Natural Rubber by Means of Infrared Spectroscopy. Part VI. Aging by Heat and Light Discussion," *Rub. Chem. Technol.*, Vol. 29, p. 1255, 1956.

10. Field, J. E., Woodford, D. E., and Gehman, J. D., "Infrared Study of Oxidation of Elastomers," *J. Polymer Sci.*, Vol. 15, p. 51, 1955.

11. D'Or, L., and Kössler, I., "An Infrared Spectrometric Study of the Oxidation of Natural Rubber," *Rub. Chem. Technol.*, Vol. 25, p. 251, 1952.

G. Creep of Elastomers Undergoing Aging,

J. Moacanin, J. J. Aklonis,[2] R. F. Fedors, and R. F. Landel

1. Introduction

A theoretical analysis of the viscoelastic behavior of elastomers undergoing scission reactions was reported in SPS 37-54, Vol. III, pp. 105–107. This initial analysis provided a conservative estimate of fatigue life. The present theoretical development takes into account the effect of sample history on the creep behavior.

2. Discussion

In general, when considering the problem of predicting the creep response of an elastomer undergoing scission reactions, one knows the viscoelastic response of a given non-degrading elastomer as a function of the crosslink density or number of effective chains ν. Also, if Plazek's equation (SPS 37-54, Vol. III) is generally valid, this response can be predicted over at least ten decades of time. From studies of degradation reactions at several temperatures, one should be able to deduce the relevant rate constants for the temperature of interest, e.g., room temperature. From the rate constants, one can then calculate the rate of change of ν. Thus, we wish to predict

$$J = J\,[\nu(t)]$$

where J is the shear creep compliance, while knowing from separate experiments how J depends on ν for non-degrading systems and how ν varies with time for a degrading system. Crosslinking networks, being more difficult to treat, will not be considered in this article. The following is the approach developed for the case in which scission takes place.

[2]Department of Chemistry, University of Southern California, Los Angeles, Calif.

For networks undergoing scission, we assume that the elastomer will exhibit, at a given time, the same response as a non-degrading elastomer of equal ν. This is illustrated in Fig. 16 for the case of a single change in crosslink density at t^* from $\nu_1 \rightarrow \nu_2$. The dashed curves represent the response for non-degrading elastomers having densities ν_1 and ν_2, respectively. The predicted "degrading" system, represented by the solid curve, follows the response of curve ν_1 until t^*. After t^*, it follows curve ν_2 translated horizontally by the amount a. The necessary physical assumption to justify this shift is that all of the load is uniformly distributed among the network chains before reaction, and that when scission occurs, the load is uniformly redistributed instantaneously among the remaining chains. In terms of a model representation, this would imply that we start with a sufficient number of Voigt viscoelastic elements in parallel to simulate the response of the network with ν_1 chains; at t^* we remove the elements corresponding to ν_1-ν_2 network chains. By restricting our treatment to a horizontal shift, we neglect the possibility of an instantaneous Hookean response when the load per element is increased.

In order to arrive at a mathematical representation of this treatment for a system undergoing continuous chemical change, we have assumed that the shear creep compliance $J(t)$ can be represented by the following equation:

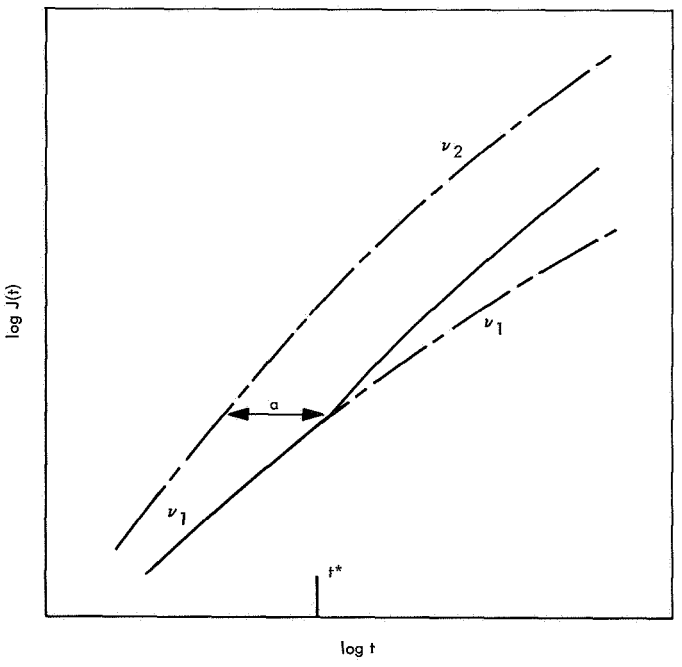


Fig. 16. Hypothetical compliance curve for a single discontinuous change in crosslink density

$$J(t) = G_g^{-1} \frac{1 + (t/\tau)^{1/n}}{1 + G_e/G_g\,(t/\tau)^{1/n}} \tag{1}$$

where G_e is the equilibrium rubbery shear modulus, G_g is the glassy shear modulus, τ is a constant, and n is an integer. Normalizing to $\tau = 1$ and defining $G_e/G_g = \epsilon$:

$$J(t) \times G_g = \frac{1 + t^{1/n}}{1 + \epsilon t^{1/n}} \tag{2}$$

Now, for a degrading system, ϵ will be a function of time (G_g is taken to be constant) and, for some experimental time t, there corresponds a shift a_t due to reactions prior to t. Then, if at t a change in ϵ (i.e., ν) occurs, we impose the condition

$$J_\epsilon(t + a) = J_{\epsilon+\Delta\epsilon}(t + a + \Delta a) \tag{3}$$

Substituting Eq. (2) and using the variable $u = t + a$

$$\frac{1 + u^{1/n}}{1 + \epsilon u^{1/n}} = \frac{1 + (u + \Delta a)^{1/n}}{1 + (\epsilon + \Delta\epsilon)(u + \Delta a)^{1/n}} \tag{4}$$

Using the binomial expansion and going to the limit of $\Delta\epsilon \to 0$, the following differential equation ensues:

$$\frac{da}{d\epsilon} = \frac{nu(u^{1/n} + 1)}{(1 - \epsilon)} \approx nu(u^{1/n} + 1) \tag{5}$$

Inasmuch as $G_g \approx 10^{10}$, ϵ will be of the order 10^{-4} and, therefore, the approximate form where ϵ is neglected relative to unity should be adequate for our purposes.

The degradation kinetics will determine the time dependence of ϵ. It turned out that for this non-linear differential equation, a separation of variables could be obtained only for zero-order kinetics, viz., for

$$\epsilon = \epsilon_0(1 - kt) \tag{6}$$

Experimental evidence for such a scission rate law has been given by P. Thirion (Ref. 1), among others. Therefore, one obtains

$$\frac{du}{dt} = 1 - n\epsilon_0 k u(1 + u^{1/n}) \tag{7}$$

which can be integrated to give

$$t = \int_0^t dt = \int_{\lambda=0}^{u} \frac{d\lambda}{1 - n\epsilon_0 k\lambda(1 + \lambda^{1/n})} \tag{8}$$

For first order kinetics, i.e., $\epsilon = \epsilon_0 \exp(-kt)$, a series solution was obtained for only $n = 1$. In other cases it is necessary to resort to numerical solutions using a computer.

Although we have presented these results as applied to the calculation of the creep behavior of networks experiencing scission, the assumptions inherent in this treatment are equally applicable to stress relaxation. Here again, the uniform distribution of stress among all existing chains both before and immediately after the scission process is assumed.

Using the reciprocal of Eq. (1) as an approximate generating function for $G(t)$, the shear stress relaxation modulus of a polymer with crosslink density ν, and following the same logic as in the case of creep, one eventually arrives at Eq. (5). Thus Eq. (5) is valid irrespective of whether stress relaxation or creep is considered. It should be noted, however, that the parameters of the generating function, n, G_g, and G_e, may be different for the *same* polymer undergoing stress relaxation or creep. (Ref. 2).

Under certain conditions, the response of a network can be dominated by either viscoelastic retardations (i.e., very slow chemical scission) or scission phenomena (i.e., very fast retardation). Since our development is designed to treat the situations where both of these effects contribute substantially to the observed behavior, it will be necessary to determine the range of values of the parameters ϵ_0 and k in which a coupling of the two mechanisms will occur.

A reexamination of Eq. (8) in this context is helpful; if retardation is fast relative to degradation (k is very small), then the second term in the denominator of the integrand will be negligible. Under these conditions we can write

$$t = \int_0^u d\lambda = u \tag{9}$$

or, since $a = u\text{-}t$, $a = 0$. This is merely the mathematical statement of the intuitively obvious fact that very slow degradation may be ignored in the considerations of the viscoelastic response on a realistic time scale. For fast degradation, on the other hand, the network will be destroyed (i.e., $\epsilon \to 0$) before appreciable creep can take place.

To analyze the behavior intermediate to these two extremes, Eq. (8) can be written as

$$t = \int_{\lambda=0}^{u} \left[1 + n\epsilon_0 k\lambda\,(1 + \lambda^{1/2})\right] d\lambda \qquad (10)$$

where we have expanded the denominator of the integrand using the condition that the second term is small compared to unity. Also, since the values of interest of the quantity u are orders of magnitude greater than 1, Eq. (10) is further simplified and integrated to yield

$$t = u\left[1 + \frac{n^2}{2n+1}\,\epsilon_0 k \exp_u\left(1 + \frac{1}{n}\right)\right] \qquad (11)$$

The coupling of the two network mechanisms is reflected in the value of a. Consider, for $n = 1$, the conditions that lead to a becoming $t/10$. Thus

$$\frac{\epsilon_0 k}{3}\,u^2 = 0.1 \qquad (12)$$

gives the value of u when $t/u = 1.1$ or, since $t \approx u$,

$$t \approx \frac{3}{10\epsilon_0 k} \qquad (13)$$

Furthermore, we can constrain this value of a to occur before half of the network has decayed, i.e., $1 - kt \geq 1/2$ or

$$u \approx t \leq \frac{1}{2}\,k \qquad (14)$$

Substitution of Eq. (14) into Eq. (13) yields the necessary relationship between k and ϵ_0

$$k \leq \frac{\epsilon_0}{1.2} \qquad (15)$$

Thus, if Eq. (15) is satisfied, the value of a will be 10% of t before half of the network has decayed. The maximum value of k that will result in a coupling of retardations and scission is $k \cong \epsilon_0$. At values of k greater than this, $\epsilon \to 0$ too rapidly. Similar results for other values of n are found in Table 8.

The predicted creep behavior for two values of k/ϵ_0 is plotted in Fig. 17. The solid line is the response of a non-degrading network as predicted by the generating function given in Eq. (1) with $\tau = 1$ s, $G_g = 10^{10}$ dyn/cm^2, and $\epsilon_0 = 10^{-4}$. The predicted responses for the corresponding reacting systems for $k = 10^{-4}$ s^{-1} and $k = 10^{-6}$ s^{-1} are shown as a dashed line and dotted line, respectively. For comparison, several values of $J(t)$ without the application of the a shift, are shown as crosses. This is the response that would be calculated if degradation were assumed to change G_e only, with no effect on the course of the retardation process.

Table 8. Some constraints on the scission-rate constant k to realize mechanical–chemical coupling

Exp n	Time when t/u exceeds 1.1	No coupling will be observed if k is greater than:
1	$\left(\frac{3}{10k\epsilon_0}\right)^{1/2}$	ϵ_0
2	$\left(\frac{5}{40k\epsilon_0}\right)^{2/3}$	$(3\epsilon_0)^2$
5	$\left(\frac{11}{250k\epsilon_0}\right)^{5/6}$	$(10\epsilon_0)^5$
10	$\left(\frac{21}{1000k\epsilon_0}\right)^{10/11}$	$(250\epsilon_0)^{10}$

It is apparent that the shift in the predicted response curves measured by the quantity a is significant in this range where both retardation and degradation are important, and that the curve for $k = 10^{-4}$ s^{-1}, ($k/\epsilon_0 = 1$, Table 8) represents the fastest degradation rate where this coupling effect can be observed.

It is to be noted that in spite of the fact that considerable reference has been made to Eq. (1), the generating function for the retardation response of non-reacting networks, the treatment embodied in this article does not hinge on the applicability of this function to real

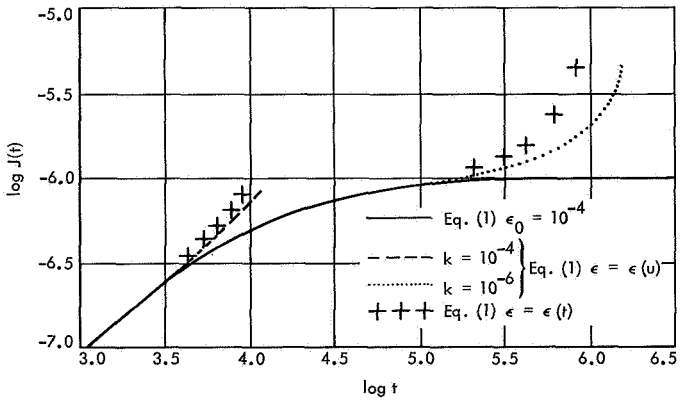


Fig. 17. Coupling of viscoelastic retardation and chain scission

systems. It was chosen merely as a matter of convenience and can be readily replaced by whatever function is appropriate to the real system of interest.

These ideas are presently being applied to an analysis of the apparent viscoelastic response of systems undergoing simultaneous relaxation and scission for which experimental results are available.

References

1. Thirion, P., and Chasset, R., *Rub. Chem. Technol.*, Vol. 37, p. 617, 1964.

2. Aklonis, J. J., and Tobolsky, A. V., *J. Appl. Phys.*, Vol. 37, p. 1949, 1966.

XVI. Research and Advanced Concepts

PROPULSION DIVISION

A. Initial Operation of the Thermionic Diode Kinetics Experiment, *H. G. Gronroos, M. L. Peelgren, and J. P. Davis*

1. Introduction

The thermionic diode kinetics experiment described in SPS 37-54,Vol. III, pp. 119–124, is in its initial operating phase. The purpose of the experiment is to study thermionic reactor powerplant dynamics and control by nonnuclear means. The experimental assembly contains the major components in such a system, i.e., thermionic diodes, a liquid-metal heat-rejection system, and a power conditioner for the electrical output from the diodes (described in SPS 37-54, Vol. III, pp. 49–50, and in *Chapter IX-E* of this SPS volume). The specially designed auxiliary equipment consists of a switching unit for selection of diode electrical coupling patterns, a low-voltage electrical load, a transformer for inducing a variable voltage for measurement of current–voltage characteristics, and an analog computer for simulation of the neutron kinetics equations. The equipment being used in the experiment allows great flexibility, as well as adequate simulation of operating conditions in a true nuclear reactor system.

2. Experimental Assembly

a. Description of equipment. A schematic of the equipment being used in the experiment was presented in SPS 37-54, Vol. III (Fig. 5, p. 121). The nominal operating conditions are given in Table 1. The basic subsystems may be grouped into four functional categories.

Diode inputs. Each of the four installed diodes (space available for two more) has its own electron-beam and cesium-reservoir heater power supplies. These units can operate directly from a controlled parameter such as temperature and current or can be slaved to an output of the analog-computer reactor simulator.

Diode outputs (heat and electricity). The heat-transfer fluid, NaK-78, is pumped around the heat-rejection loop shown in Fig. 1. The principal components are a 25-kW

Table 1. Nominal operating conditions

Parameter	Value
NaK heat-rejection system	
Temperature, °F	1200
Flow rate, gal/min	10
Pressure drop, psi	12
Thermionic diodes (30-cm² area each)	
Maximum emitter temperature, °C	1850
Average emitter temperature, °C	1800
Collector temperature, °C	~700
Nominal operating point, V	0.6[a]
Electrical power output, W	250
Thermal power input, W	2100
Efficiency at terminals, %	12

[a]At 420 A.

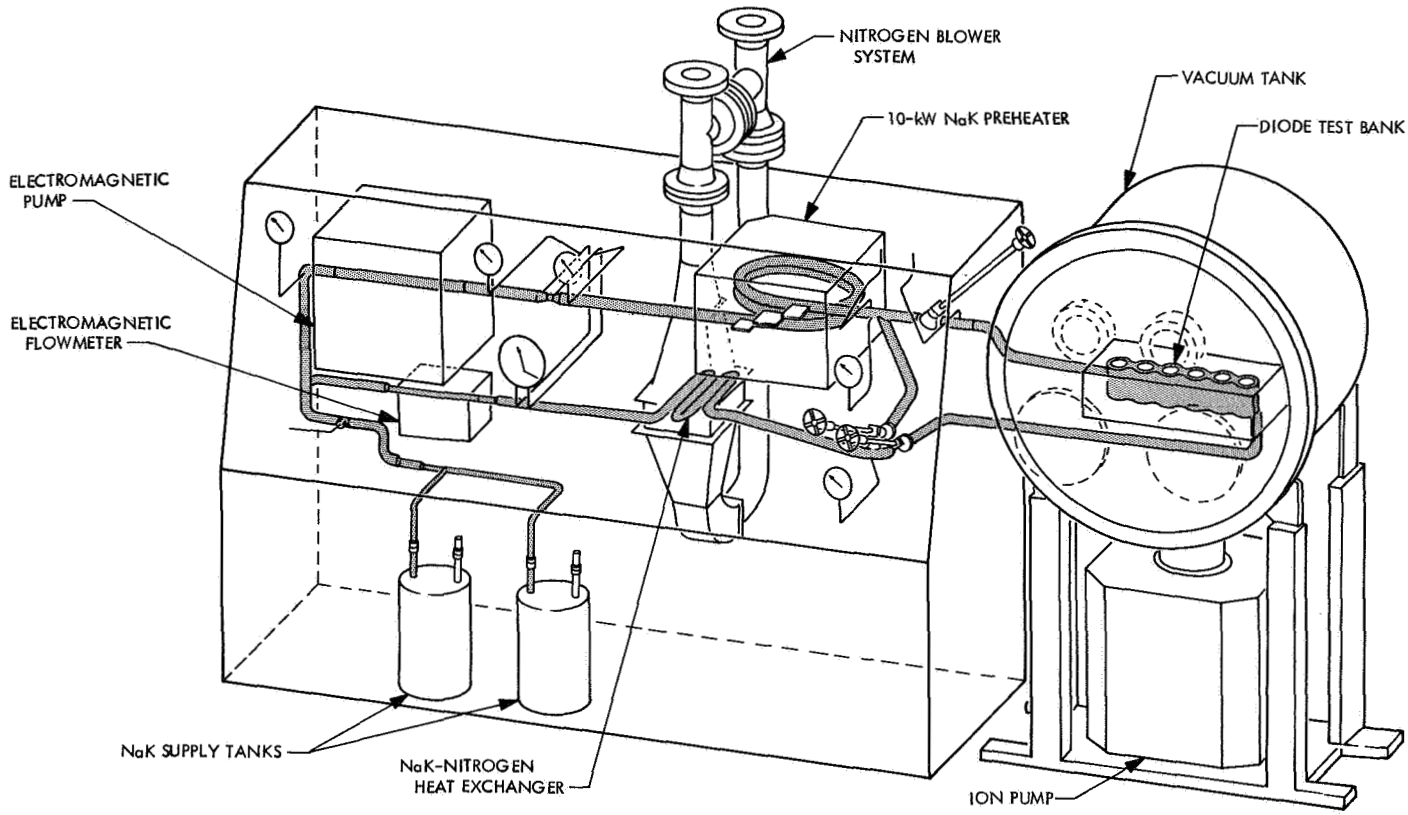


Fig. 1. NaK loop for thermionic diode kinetics experiment

NaK–nitrogen heat exchanger, an electromagnetic pump, two flowmeters (one venturi and one electromagnetic) and a 10-kW preheater. The nitrogen from the heat exchanger is recirculated after being cooled by an air–nitrogen heat exchanger. The 10-kW preheater is the only component in the heat-rejection system that is either controlled by temperature or programmed from the analog computer.

The electrical output from the diodes goes through current shunts to a coupling unit. This unit consists of an array of 22 600-A mercury switches for all interconnections to the output buses and a control panel for the activation, monitoring, and safeguarding of the switches. From the coupling unit output, the power can be put into either a low-voltage load or a power conditioner for stepping up to high voltage. The electrical output circuit for the diodes is shown in Fig. 2. The power conditioner used in this experiment is that described in *Chapter IX-E.*

The low-voltage load consists of a water-cooled hairpin loop (with insulated flow return bend), a variable mercury column external to the hairpin tube, and a mercury-height-adjusting cylinder. The stainless-steel hairpin consists of two thin-walled tubes, approximately 36 in. long, connected by a phenolic U-bend. This assembly is brazed to copper mounting pads at the top and is installed in a 1-in.-ID glass tube. This tube, connected at the bottom to a mercury reservoir, is filled to any desired height corresponding to the desired load impedance. The mercury level is adjusted manually by means of a handle-and-screw arrangement on a double-acting Belofram cylinder. When the load is in the open-circuit position (when the power conditioner is in use), the cylinder can be pressurized from a fault-indicating signal from the power conditioner. This pressurization results in the breakaway of the magnetic latch and displaces the mercury to a predetermined level. This feature prevents the diodes from running at open-circuit conditions under excess power due to a malfunction of the power conditioner.

A sweep transformer, used in conjunction with the low-voltage load, induces an ac voltage on the diode output to obtain a current–voltage characteristic of the diode. The initial stepdown from 480 V is through an autotransformer. This output is put into the primary coils

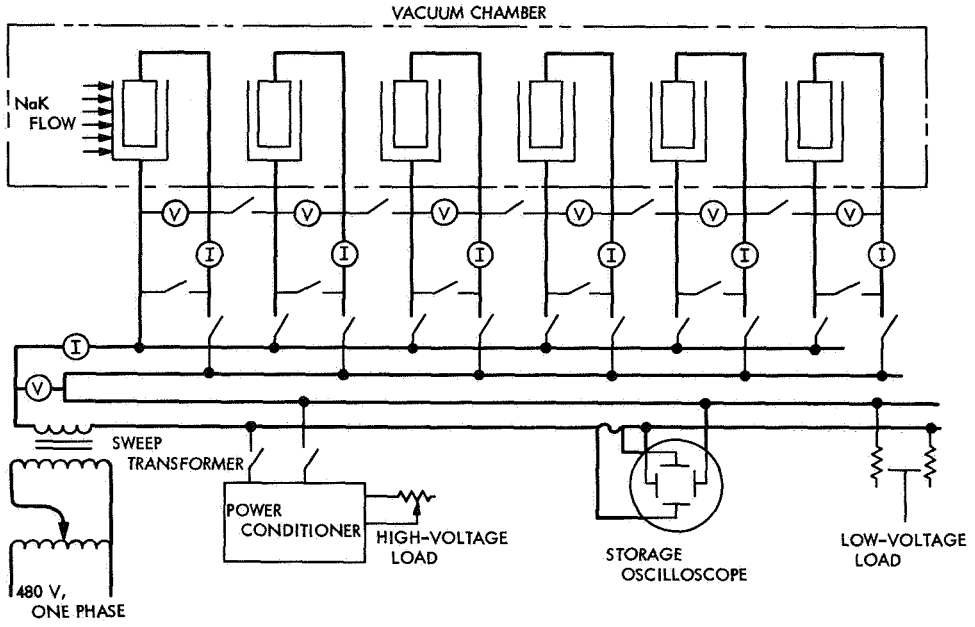


Fig. 2. Diode electrical output circuit

of the sweep transformer. The secondary of the sweep transformer is two turns of 1- × 4-in. copper bus. The sweep transformer has a capability of 12 V and 10^4 A.

Data acquisition and controls. Dymec data acquisition and printout are used along with numerous display meters and 24 channels of chart recorders.

Auxiliary equipment. The auxiliary equipment necessary for safe and efficient operation includes an extensive array of safety interlocks; vacuum, argon, nitrogen, and air systems; and several water- and oil-cooling systems.

b. Problem areas. The few problems encountered thus far have been minor and either have random causes (e.g., power outages or trips of vacuum pumps or other equipment essential to conduct of the test), or are intrinsically related to the test itself. Among the latter type of problems, the mercury switches have malfunctioned through leakage of: (1) the top O-ring seal, and (2) the Belofram diaphragm. All leaks were eliminated by reworking the units. The second area of some concern is the W–5% Re and W–26% Re thermocouples. There have been several mechanical failures of the brittle wire at the top of the sheath, as well as shorting of the thermocouple to its tantalum sheath, which is at emitter potential. The third problem area is the shorting of the filaments due to misalignment or sag of the filament in the emitter. These problems have been corrected and have had little influence on the test program.

3. Thermionic Diode Performance

a. Temperature distributions. All four thermionic diodes have been performance-tested. Data were compared with the results obtained by the manufacturer (Thermo Electron Corp., Waltham, Mass.) using a special water-cooled test rig. An absolute comparison is not possible, since the exact collector temperature is not available; however, its range is 600–700°C. In the NaK coolant loop, the collector temperature can be varied by adjusting the inlet temperature to the test bank and by adjusting the helium pressure in the gap between the thimble and the diode structure. However, the data indicate that the diode performance is essentially the same in the NaK coolant system and in the water-cooled test rig.

The emitter thermocouples are inserted near the bottom end, in the middle, and near the top end of the emitter length. The thermocouples measure a temperature higher than the emitter surface temperature. Calculations using a computer code for three-dimensional heat transfer show that, in the range of nominal operating conditions, the temperature drop across the emitter wall is about 100°C, and the thermocouples read about 50°C over the emitter surface temperature. Although accurate diagnostic measurements are not possible, the available

thermocouples are very useful for operation of the experimental assembly and for investigation of the dynamic behavior. Also, the combination of the thermocouples and the above-mentioned computer program enables adequate mapping of the temperature distribution.

The electron beam filaments consist of coiled tungsten wire. In the original design, the coil was wound with about a 1:4 pitch variation from bottom to top in an attempt to flatten the temperature distribution. This design overcompensated and was thus unsatisfactory. An evenly wound coil with a single extra turn at the top and bottom ends was also used. This coil minimized the temperature drop at the bottom and gave a distribution that more nearly approximates that existing in a nuclear reactor. However, this coil appears susceptible to a temperature run-away phenomenon where the temperature at the bottom end increases sharply, the middle remains about the same, and the top decreases sharply. This phenomenon occurred at the lower cesium reservoir temperatures and appeared to be connected with "thermionic burn-out" (discussed in *Paragraph c*). The final filament coil design consists of an evenly wound coil with an extra turn at the top end only. These filaments will be installed at the next opportune time.

b. Current–voltage characteristics. Figure 3 shows current–voltage characteristics obtained by inducing a 60-Hz sinusoidal voltage in the thermionic diode circuit and displaying the resulting trace on a memory oscilloscope. As indicated on the figure, this measurement technique does not give exactly the same characteristics as those obtained by static measurements due to the shift in the axial temperature distribution with the static operating point. The oscilloscope traces are produced with a phase-shift-correcting network that eliminates the loop trace that would otherwise result due to distributed circuit inductances. Also, the operating point that was selected ensures minimum broadening of the traces by harmonics. A final precaution is to maintain the induced voltage for only a fraction of 1 s and thus limit the emitter surface temperature fluctuation to an estimated value of less than 10°C. The oscilloscope traces show the "instantaneous" characteristic for a step change in electric load and give a good approximation of the static characteristic. Data from the oscilloscope traces are readily obtainable, while measurement of the static curve is time-consuming.

c. Thermionic burn-out. Thermionic burn-out occurs when an increase in heat input to the emitter, at cesium pressure at or below optimum, increases the emitter

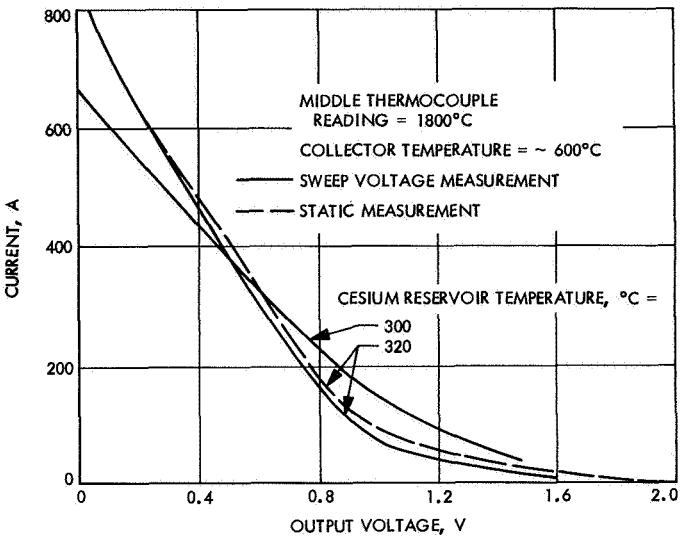


Fig. 3. Current–voltage characteristics obtained by inducing a 60-Hz sinusoidal voltage

temperature and causes the cesium to desorb from the emitter surface. The emitter temperature increases further due to decreasing electron cooling, which further desorbs cesium, etc., resulting in a runaway condition. The final equilibrium attained has a bare emitter at very high temperature, with essentially no electrical power output.

Thermionic burn-out and its consequences on reactor stability have been analyzed at Fairchild-Hiller (Ref. 1) and JPL (Ref. 2). The results indicate that self-sustained oscillations can occur in a nuclear reactor application. However, by proper control of the cesium reservoir temperature, this situation can be avoided.

To the author's knowledge, there have been no previously reported experimental observations of thermionic burn-out. However, a clear indication of its existence was obtained in this experiment. One technique was by optimizing the cesium reservoir temperature at a lower emitter temperature and then increasing the thermal power input in small steps until the diode failed to reach a new equilibrium. The burn-out is slow for this case, because it starts in a localized area of the emitter and gradually spreads over the whole surface. The second technique, the more rapid means by which burn-out was induced, was by lowering the cesium reservoir temperature. The resulting traces for the two cases are shown in Fig. 4. The thermal power input was reduced at high emitter temperatures during the transient to avoid damaging the diode.

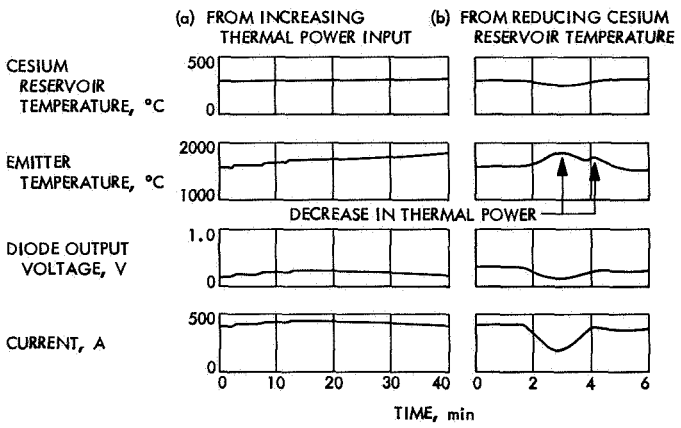


Fig. 4. Thermionic burn-out experimental results

4. System Experiments

Several experiments have been made to study the effects of: (1) coupling diodes in series and parallel, (2) perturbing one of the diodes in the chosen coupling pattern, and (3) open and short circuiting. These experiments were conducted to familiarize personnel with the system behavior. During the initial runs, the analog-computer reactor simulator was used on a limited scale on-line.

5. Concluding Remarks

During the initial operating phase of the thermionic diode kinetics experiment, the equipment operated very satisfactorily with only minor problems occurring. The assembly has been developed to a point where routine operations and procedures have been established and useful experimental information is being obtained. The next phase will be the systematic execution of the experimental program.

References

1. Schock, A., "Effects of Cesium Pressure on Thermionic Stability," paper presented at the Second International Conference on Thermionic Electrical Power Generation, Stresa, Italy, May 27–31, 1968.

2. Gronroos, H., and Davis, J. P., "Stability and Control Considerations for Thermionic Reactors," paper presented at the Second International Conference on Thermionic Electrical Power Generation, Stresa, Italy, May 27–31, 1968.

B. Liquid-Metal MHD Power Conversion: Impinging Nozzle Experiment, *L. G. Hays*

Liquid-metal magnetohydrodynamic (MHD) power conversion is being investigated as a power source for nuclear-electric propulsion. A liquid-metal MHD system has no moving mechanical parts and operates at heat-source temperatures between 1200 and 1400°K. Thus, the system has the potential of high reliability and long lifetime using readily available containment materials such as Nb–1%Zr.

Fabrication of a 50-kWe NaK–nitrogen conversion system and a 100-kWt cesium–lithium loop is continuing. Tests of a new type of separator with no wall friction are described below.

Friction on the surface-impingement type separator is responsible for a major loss in the MHD system. Analysis has shown that the cycle efficiency would increase by two percentage points if the friction were eliminated. One approach to friction reduction is to use two nozzles for flow acceleration and to mix the exit jets by impingement on each other. In this way, it might be possible to sufficiently concentrate or coalesce the liquid phase without the introduction of a solid surface.

To test this concept the apparatus shown in Fig. 5 was fabricated for operation with water–nitrogen mixtures. Water and nitrogen are mixed by the injector assembly and are accelerated to 70–140 m/s in the two-phase nozzles. The angle of impingement between the two nozzles can be varied. The exit jets impinge in the mixing region, where they can be studied visually and by probes. Different portions of the resultant flow are captured by the adjustable knife edges. The flow rate and thrust produced by the captured portion of the flow are measured to determine the degree of concentration of the liquid.

The initial experiments have shown that a concentration of the liquid does occur. Figure 6 shows the mixing region with the two nozzles inclined 15 deg to the axis and with a nozzle exit velocity of 85 m/s. The results of a survey of the stagnation pressure at the location shown are presented in Fig. 7. Approximately 90% of the 65-kg/s liquid flow is concentrated in a region 2.0 cm wide. This represents a decrease in the gas/liquid volume ratio from 11.6, the value at the nozzle exit, to about 4. A value of 4 can be accommodated by the MHD generator in a cesium–lithium system by dissolving the vapor in the upstream diffuser. Therefore, if this value of volume ratio can be attained for the cesium–lithium system, the surface-impingement separator can be eliminated. Experiments are continuing over a range of nozzle inlet pressures, mixture ratios, and impingement angles.

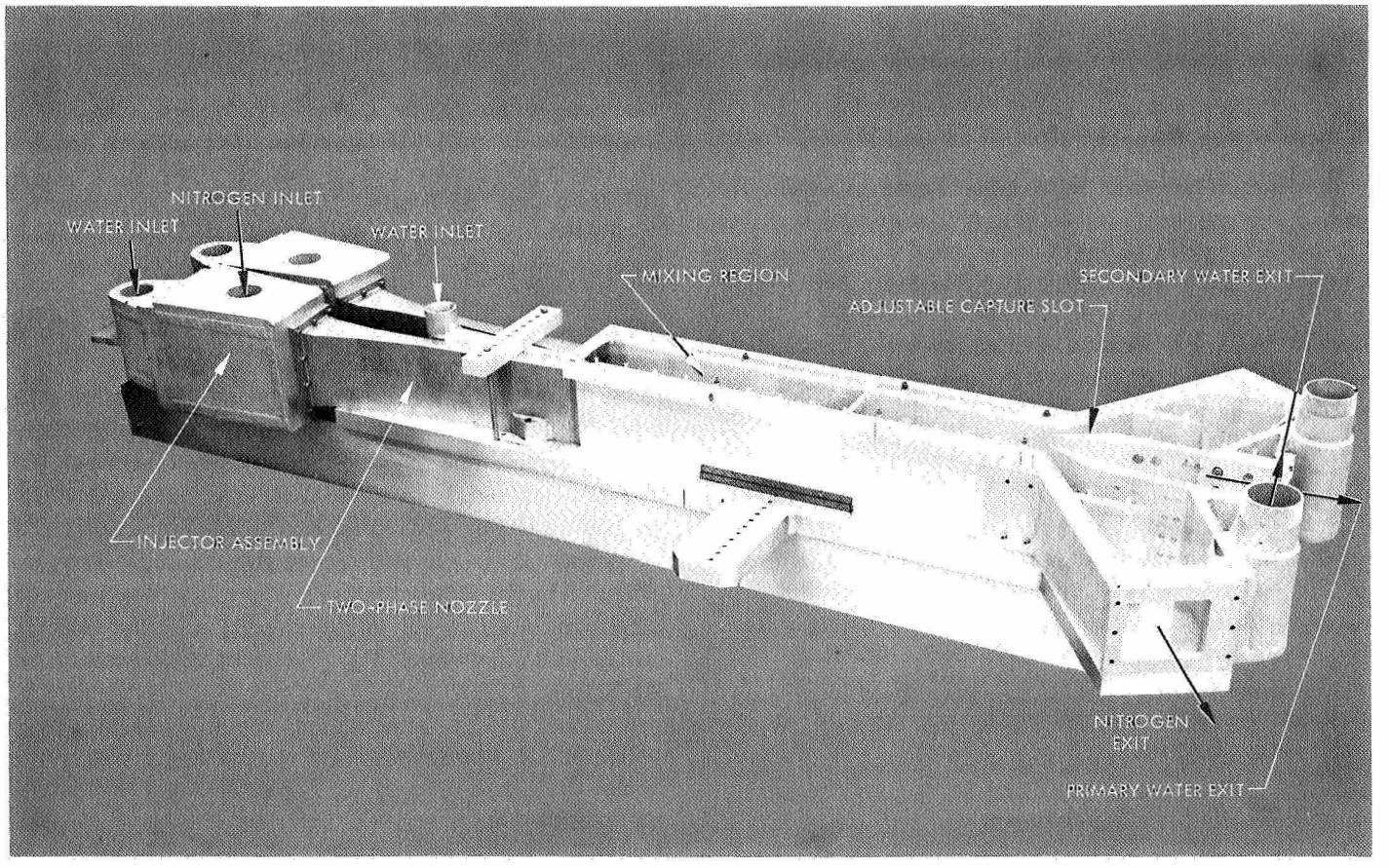


Fig. 5. Impinging nozzle test assembly

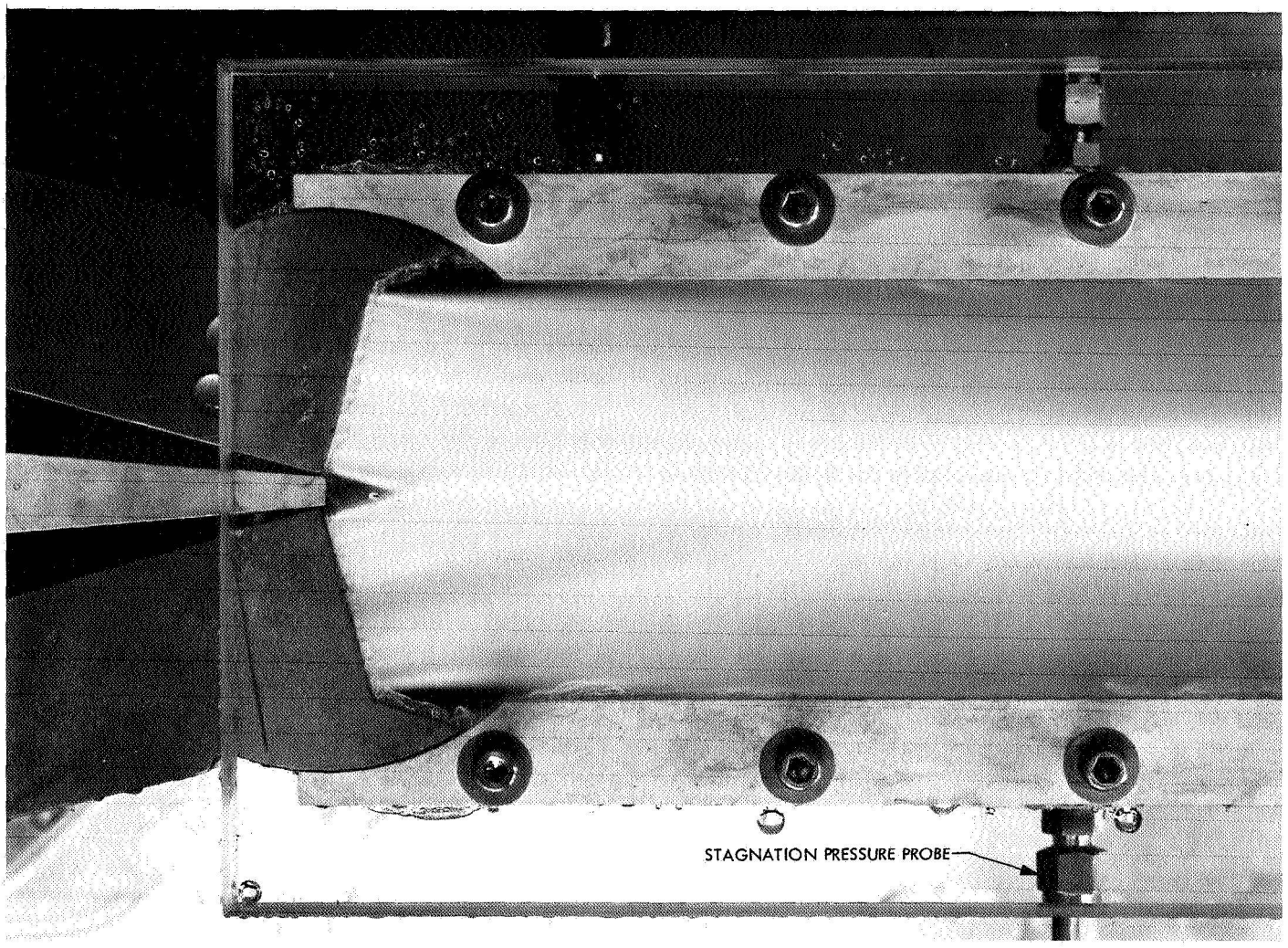


Fig. 6. Impinging nozzles in operation

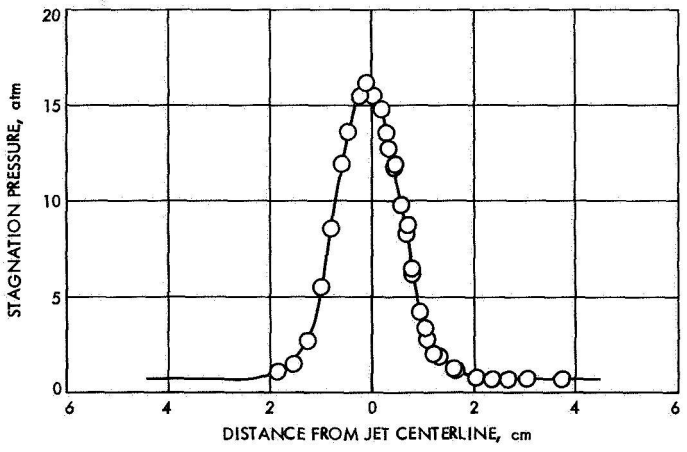


Fig. 7. Stagnation pressure profile across coalesced jet

C. Hollow-Cathode Operation in the SE-20C Ion Thruster, *T. D. Masek and J. R. Womack*

Many of the possible trade-offs in the use of hollow cathodes in ion thrusters have been discussed previously.[1] The present studies are intended to provide additional data on the operation of a hollow-cathode thruster with certain pole piece and baffle configurations. These data will be useful in choosing a suitable design for the SEPST (solar electric propulsion system technology) III program. The SE-20C thruster used in these tests was described in SPS 37-51, Vol. III, pp. 124–128.

Tests were conducted with a single pole piece and four baffle configurations. The basic pole piece and baffle configuration is shown in Fig. 8. Table 2 lists each baffle size used. The cathode, also used in the testing discussed in Ref. 3, had an orifice diameter of 0.063 cm and a tip diameter of 0.42 cm. The cathode was fitted tightly in the boron nitride mount, with the tip flush with the boron nitride surface.

Data on discharge power per beam ion, accelerator impingement, and thruster stability are of particular interest, since these factors largely determine the usefulness of a given design. For the baffles tested, the discharge power per beam ion did not vary greatly with baffle size. Typical discharge chamber losses (ion chamber power/ion beam current) are plotted in Fig. 9 as functions of mass utilization efficiency (mercury flow in equivalent amperes/ion beam current) for configuration 4. For all tests, the discharge losses showed a substantial dependence upon the flow rate through the cathode.

Accelerator impingement was also found to be relatively independent of baffle size. These data are shown in Fig. 10. The most important feature of these data is the sharp rise in impingement for high utilization efficiencies and high total flow rates. This effect, not observed with the oxide cathode in the normal operating range, appears to be due to the "direct propellant introduction" from the cathode (Ref. 4). A portion of this flow can reach the grids un-ionized. To prevent this direct loss of propellant and thus achieve a given utilization efficiency, the density of the center portion

[1]SPS 37-48, Vol. III, pp. 119–125; SPS 37-49, Vol. III, pp. 207–211; and Refs. 1–3.

Table 2. Baffle and diffuser configurations

Configuration	Diameter, cm		Material	
	Baffle	Diffuser	Baffle	Diffuser
1	2.79	—	Stainless steel	—
2	3.05	—	Stainless steel	—
3	3.05	—	Ferromagnetic	—
4	3.05	3.56	Ferromagnetic	Stainless steel

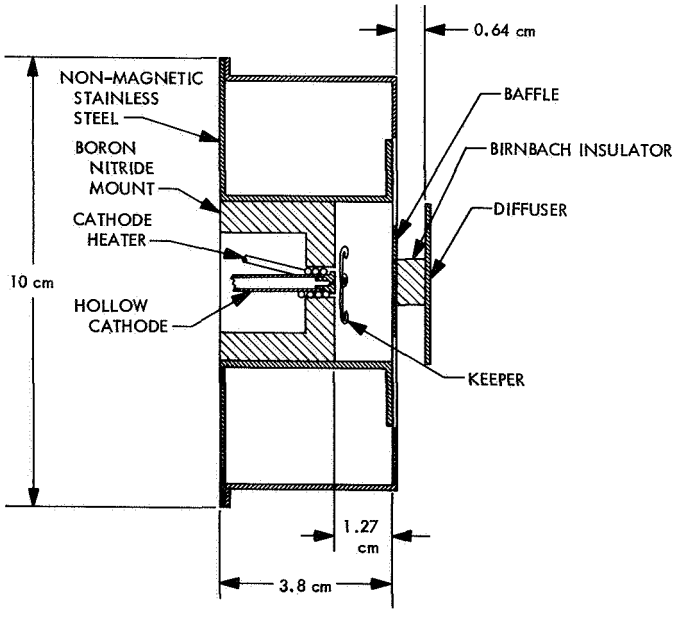


Fig. 8. Basic hollow-cathode pole piece and baffle configuration

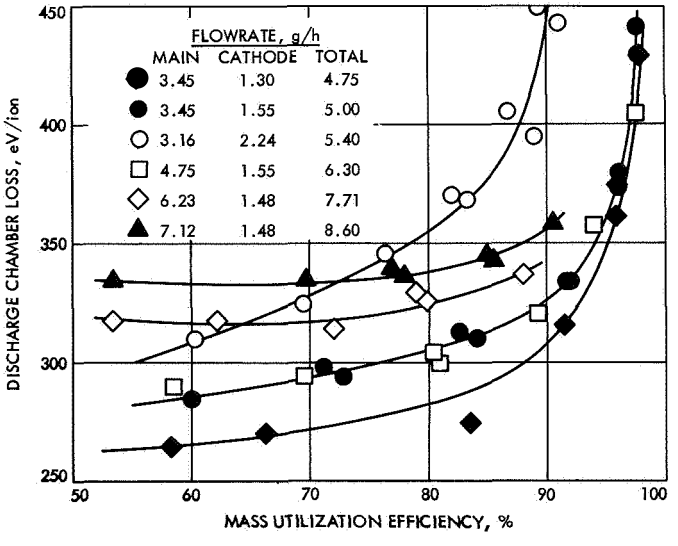


Fig. 9. Discharge chamber loss vs mass utilization efficiency for baffle configuration 4

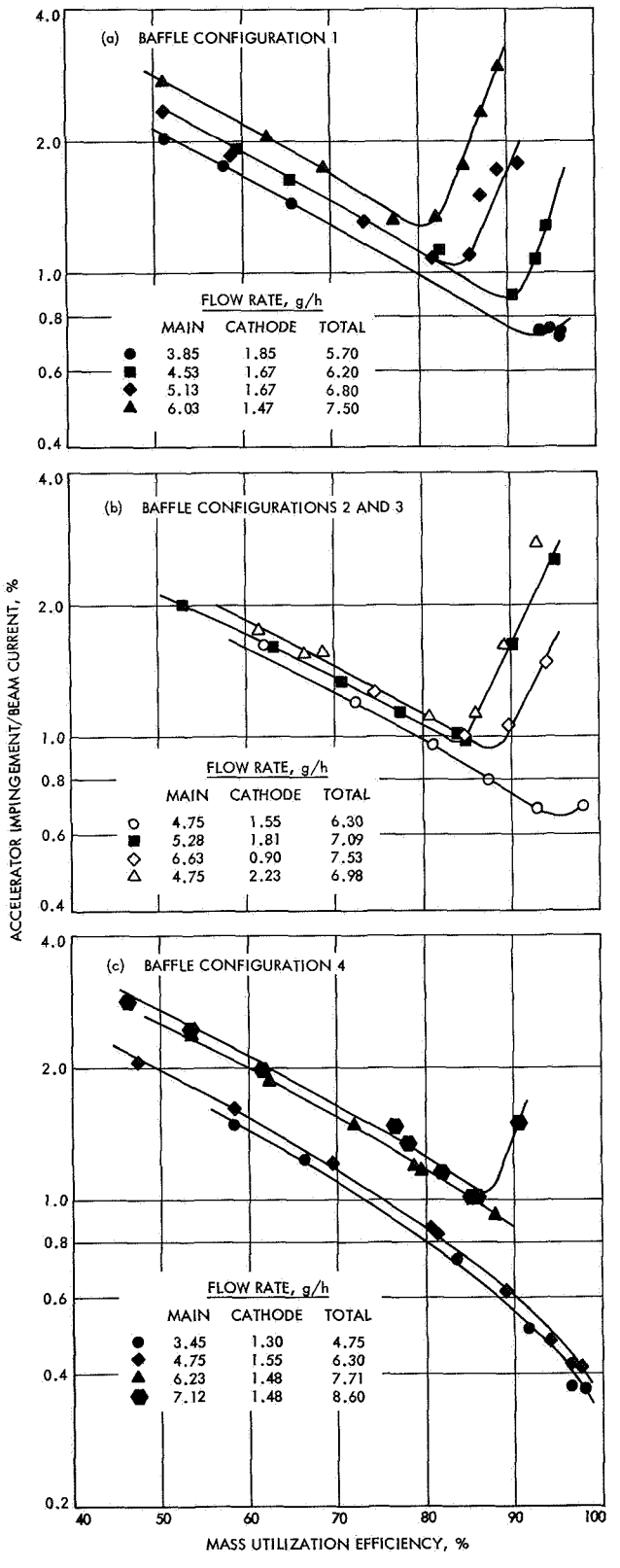


Fig. 10. Accelerator impingement/beam current vs mass utilization efficiency

of the discharge plasma should be increased compared with that for no cathode flow (oxide cathode).

The ion acceleration optics are limited to a maximum current density. When this maximum is exceeded, the individual ion beams become unfocused and impinge on the accelerator grid. The observed rise in impingement (Fig. 10) appears to be due to this effect.

The diffuser (Fig. 8) was added to the baffle to reduce the direct loss of propellant from the hollow cathode. This element allowed higher beam currents to be drawn while maintaining "normal" impingement (950 mA, compared with about 800 mA without the diffuser). Although the impingement rise was not completely eliminated by the diffuser, use of this type of element was shown to be quite beneficial.

References

1. Pawlik, E. V., *Power Matching of an Ion Thruster to Solar Cell Power Output*, Technical Memorandum 33-392. Jet Propulsion Laboratory, Pasadena, Calif., July 15, 1968.
2. Masek, T. D., and Pawlik, E. V., "Thrust System Technology for Solar Electric Propulsion," Paper 68-541, presented at the AIAA Fourth Propulsion Joint Specialists Conference, Cleveland, Ohio, June 1968.
3. Rawlin, V. K., and Kerslake, W. R., "Durability of the SERT II Hollow Cathode and Future Applications of Hollow Cathodes," Paper 69-304, presented at the AIAA Seventh Electric Propulsion Conference, Williamsburg, Va., Mar. 1969.
4. Masek, T. D., "Plasma Properties and Performance of Mercury Ion Thrusters," Paper 69-256, presented at the AIAA Seventh Electric Propulsion Conference, Williamsburg, Va., Mar. 1969.

D. Laminarization of a Turbulent Boundary Layer in a Nozzle With Wall Cooling, *R. F. Cuffel, L. H. Back, and P. F. Massier*

1. Introduction

The effect of laminarization of an initially turbulent boundary layer on heat transfer has been observed under certain flow acceleration conditions in supersonic nozzles (e.g., Refs. 1 and 2). A consequence of the laminarization phenomenon is a reduction in the local convective heat transfer, which is of importance in the design of rocket nozzles. Presented herein for the conditions at which laminarization occurred are experimental distributions of the velocity and temperature in the boundary layer at one axial location in a nozzle having 10-deg half angles of convergence and divergence. The boundary-layer probing station, shown in Fig. 11, was located

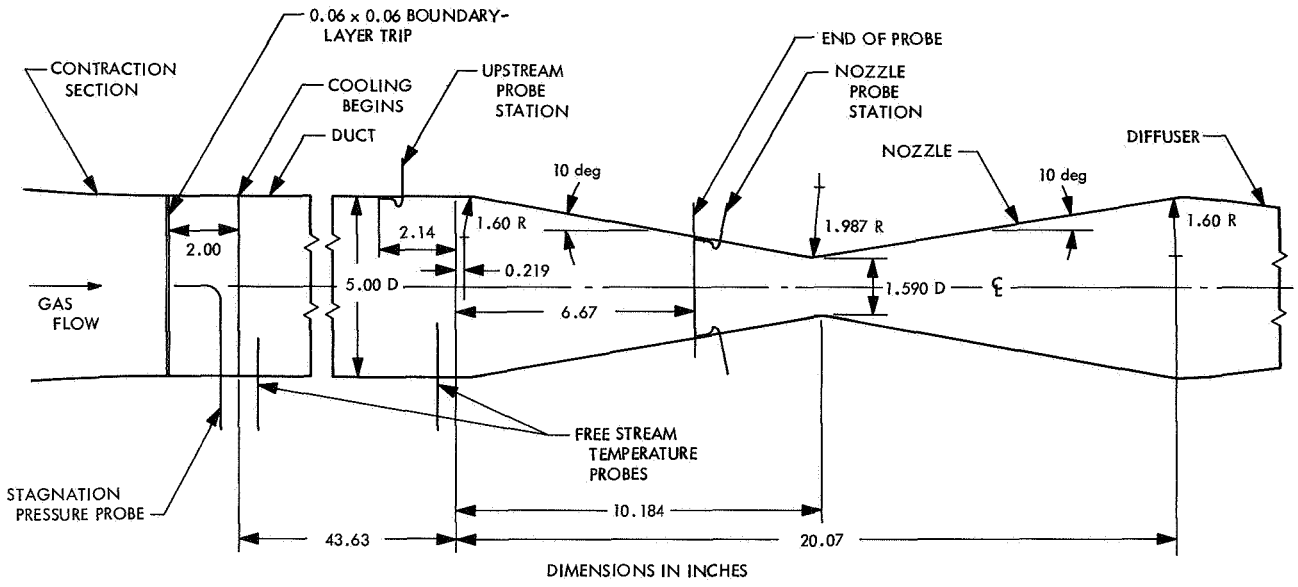


Fig. 11. Schematic showing probe location

where the greatest reduction in heat transfer (about 50%) was observed. There the Mach number was about 0.2. Laminarization was observed to occur when the laminarization parameter K was larger than about $2\text{–}3 \times 10^6$. K is defined as

$$K = \frac{\mu_e}{\rho_e u_e^2} \frac{du_e}{dx}$$

with μ = viscosity, ρ = density, and u = velocity, with e denoting conditions at the free-stream edge of the boundary layer, and with x = distance along the wall in the flow direction.

Compressed air, heated at a remote distance upstream by the internal combustion of methanol, flowed through a cooled duct before entering the cooled nozzle. Heat-transfer measurements were made in numerous coolant passages by calorimetry. Some of these measurements were previously reported in SPS 37-49, Vol. III, pp. 212–213, and SPS 37-50, Vol. III, pp. 186–189.

The nozzle boundary-layer measurements were made with a pitot probe that had an 0.005-in.-high flattened tip and with an 0.010-in.-high thermocouple probe. Similar measurements for adiabatic flow were discussed in SPS 37-51, Vol. III, pp. 116–120.

2. Results

Local heat-transfer coefficients along the nozzle, when viewed in a representation found to correlate heat-transfer data for turbulent boundary layers, indicated that a reduction in heat transfer occurred along the convergent section at pressures lower than about 45 psia. For choked nozzle flow, decreasing the stagnation pressure p_t increases the value of K, since locally $K \propto 1/p_t$. Profiles of the velocity u and temperature T in the nozzle at an axial distance from the nozzle inlet of $z = 6.67$ in., where a reduction in heat transfer was observed, are shown in Fig. 12 in terms of the dimensionless parameters u^+, T^+, and y^+. These quantities are defined as follows:

$$u^+ = \frac{u}{u_\tau}$$

$$T^+ = \frac{T_w - T}{\left(\frac{q_w}{\rho_w u_\tau c_{p_w}}\right)}$$

$$y^+ = \frac{\rho_w u_\tau y}{\mu_w}$$

where w denotes conditions at the wall and

u_τ = friction velocity
$= [\tau_w/\rho_w]^{1/2}$
τ = shear stress
q_w = heat flux from the wall
c_p = specific heat at constant pressure
y = distance normal to wall

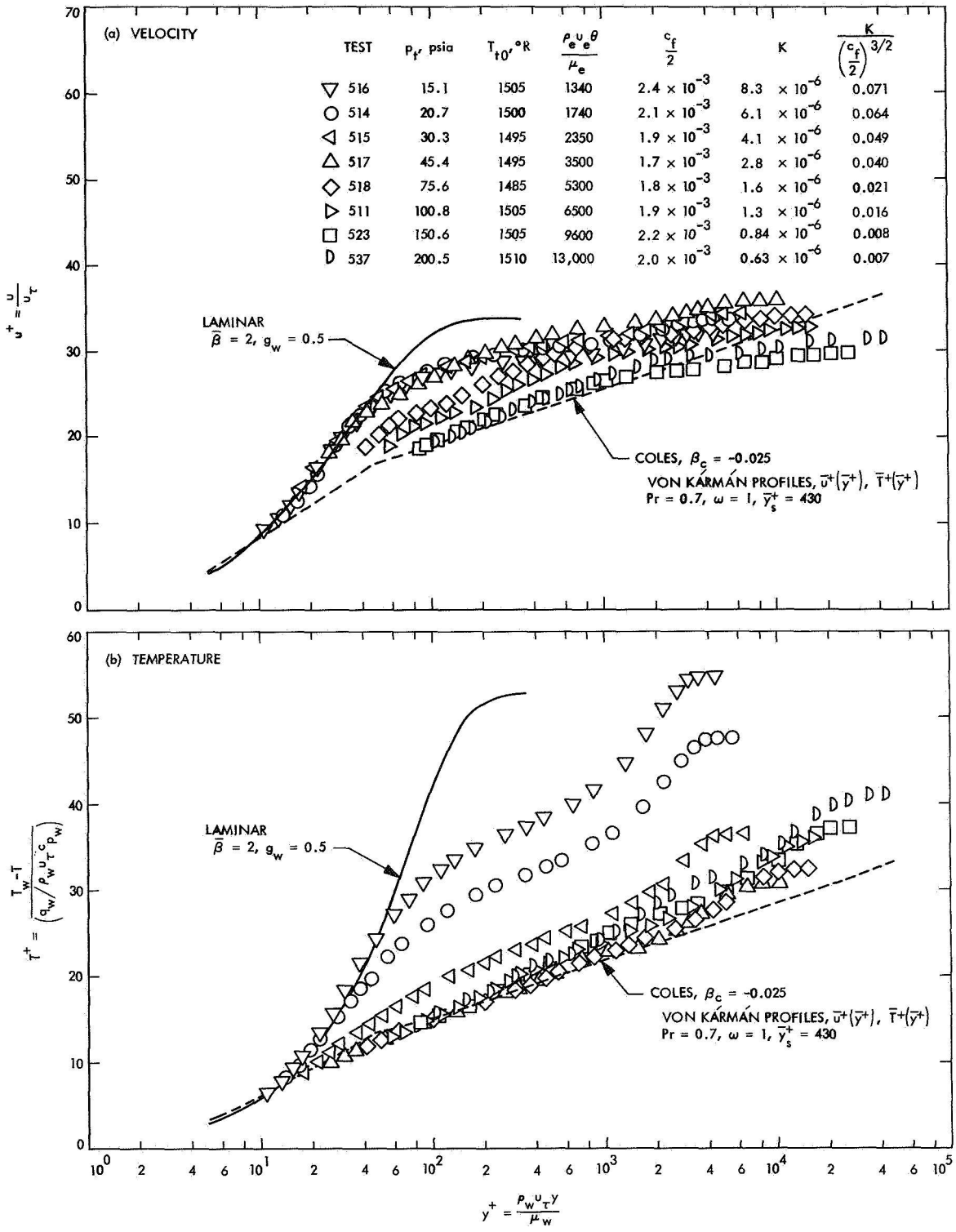


Fig. 12. Reverse transition velocity and temperature profiles in the nozzle

In Fig. 12, the subscript "0" denotes a reservoir condition, the subscript "t" denotes stagnation condition, and

c_f = friction coefficient, $c_f/2 = \tau_w/\rho_e u_e^2$

θ = momentum thickness

where

$$\theta\left(r - \frac{\theta \cos 10\ \text{deg}}{2}\right) = \int_0^\infty \left[\frac{\rho u}{\rho_e u_e}\left(1 - \frac{u}{u_e}\right)\right] \times (r - y \cos 10\ \text{deg})dy$$

with r representing the nozzle radius.

In the outer region of the boundary layer, the velocity profiles are relatively flat as a consequence of flow acceleration. The wake-like behavior found upstream of the nozzle has disappeared. The temperature profiles at the higher pressures are similar to those profiles upstream of the nozzle. The wake-like behavior observed upstream is still evident. Closer to the wall, there is a region where the velocity profiles at the higher pressures exhibit a nearly semi-logarithmic behavior found in turbulent boundary layers. At the lower pressures, the velocity profile in the wall vicinity becomes typical of a laminar boundary-layer profile. Therefore, the friction coefficients for the lower-pressure tests in Fig. 12 were obtained by matching the velocity profiles to a laminar boundary-layer prediction with acceleration and wall cooling.[2] The particular prediction used was for the acceleration parameter $\bar{\beta} = 2$, where

$$\bar{\beta} = \frac{T_{t0}}{T_e}\frac{2\xi}{u_e}\frac{du_e}{dx}$$

ξ = coordinate in x-direction

$$= \int_0^x \rho_e \mu_e u_e r^2 dx$$

with $g_w = T_w/T_e = 0.5$ and Prandtl number $\text{Pr} = 0.7$, which closely match the experimental conditions. Viscosity was assumed proportional to temperature to the 0.7 power, ω. At the higher pressures, the friction coefficients were obtained by using as a guide the predicted velocity profiles from Cole's transformation theory (Ref. 3) for an acceleration parameter, $\beta_c = q_w/(T_w \rho_w c_{p_w} u_\tau)$, value of -0.025. The friction coefficients for the intermediate tests were obtained by using these two limiting profiles as guides and observing the slope of the measured velocity profiles in the wall vicinity.

The correspondence between the measured velocity and temperature profiles near the wall at the lower pressures and the laminar boundary-layer prediction indicates that turbulent transport could not have been significant in the wall vicinity where the production of turbulence is largest for a normal turbulent boundary layer. However, the shape of the profiles in the outer part of the boundary layer still indicates an influence of turbulent transport. Laminarization in accelerating flows apparently results from a suppression of the mechanism by which turbulence is produced in the wall vicinity. This view is consistent with other studies (Refs. 4 and 5) that indicate the importance of the region near the wall in the production of turbulence.

On closer inspection of the profiles, it is noted that the velocity profiles are typical of the laminarized behavior up to a pressure of 45 psia, while the temperature profiles at 45 psia and above collapse onto a turbulent profile. The latter behavior was also noted in the heat-transfer data where the reduction in heat transfer occurred at pressures lower than about 45 psia. The velocity profile was turbulent above about 150 psia. Based on these observations, it is apparent that, with wall cooling, the laminarizing effect of acceleration on the velocity profile can occur at Reynolds numbers larger than those at which laminarization influences the temperature profile.

3. Summarizing Remarks and Conclusion

Heat-transfer and boundary-layer measurements in the convergent section of a cooled, conical nozzle were found to be unlike those for a turbulent boundary layer when values of K exceeded about $2\text{–}3 \times 10^{-6}$.

At the larger values of K, or at lower Reynolds numbers where the reduction in heat transfer amounted to about 50%, boundary-layer measurements indicated the drastic change in the velocity and temperature profiles in the wall vicinity. The profiles became laminar-like near the wall. The mechanism by which turbulence is produced in the wall vicinity is apparently suppressed in accelerating flows. This conclusion is consistent with observations on the important role of the wall in the production of turbulence.

[2]Back, L. H., *Acceleration and Cooling Effects in Laminar Boundary Layers—Subsonic, Transonic, and Supersonic Speeds* (to be published in the *AIAA J.*).

At the intermediate values of K, the velocity profile still became laminar-like near the wall, while the temperature profile, like the heat transfer, was typical of a turbulent boundary layer.

References

1. Back, L. H., Massier, P. F., and Cuffel, R. F., "Flow Phenomena and Convective Heat Transfer in a Conical Supersonic Nozzle," *J. Spacecraft Rockets*, Vol. 4, No. 8, pp. 1040–1047, Aug. 1967.

2. Boldman, D. R., Schmidt, J. F., and Gallagher, A. K., *Laminarization of a Turbulent Boundary Layer as Observed from Heat-Transfer and Boundary-Layer Measurements in Conical Nozzles*, NASA TN D-4788. National Aeronautics and Space Administration, Washington, D.C., Sep. 1968.

3. Coles, D., "The Turbulent Boundary Layer in a Compressible Fluid," *Phys. Fluids*, Vol. 7, No. 9, pp. 1403–1423, Sept. 1964.

4. Kline, S. J., Reynolds, W. C., Schraub, F. A., and Runstadler, P. W., "The Structure of Turbulent Boundary Layers," *J. Fluid Mech.*, Vol. 30, Part 4, pp. 741–773, Dec. 1967.

5. Kim, H. T., Kline, S. J., and Reynolds, W. C., "An Experimental Study of Turbulence Production Near a Smooth Wall in a Turbulent Boundary Layer with Zero Pressure Gradient," Report MD-20. Thermosciences Div., Mechanical Engineering Dept., Stanford University, Palo Alto, Calif., Jan. 1968.

E. Reference Mission Specification[3] for 300-kWe Thermionic Spacecraft Design Study, *N. K. Simon*

Preliminary spacecraft designs are being generated using 300-kWe thermionic reactors as power sources for electric propulsion. A Jupiter orbiter mission was selected as a representative target mission so that mission-dependent parameters could be factored into spacecraft designs at these very early stages. The mission/system interface is flexible to some degree, and, in any final spacecraft design, there are tradeoffs between the mission goals and the powerplant capabilities. Therefore, in any discussion of either the mission or the system, it is necessary to make some assumptions about the other. In this case, the constraints were the following:

(1) Powerplant operation for first-generation designs will be limited to 10,000–15,000 h.

(2) The net spacecraft weight (defined as the weight of the scientific instruments, the communications and guidance equipment, and their supporting structure) will be 1 metric ton (2205 lb).

(3) The launch vehicle will be the largest of the *Titan* family, the *Titan IIIC-7*. This, as presently conceived, is a non-manrated vehicle that employs the standard trans-stage, the stretched first-stage tanks and twin 7-segment solid strap-ons.

Previous mission studies (Ref. 1) were based on an electric propulsion spiral out of earth orbit and a subsequent spiral in to the target planet's orbit. Since these missions appeared suitable for a spacecraft operating with a 300-kWe powerplant, they were used in the development of the reference mission profile. Besides the three constraints listed previously, it was necessary to make the following additional assumptions to define the reference Jupiter orbiter mission:

(1) A spiral out of earth orbit from 750 nmi.

(2) Optimum coast, best encounter.

(3) A spiral in to Jupiter orbit at 2×10^6 km.

(4) 30,000-lb initial weight for the *Titan IIIC-7* in earth orbit.

(5) 10,000-lb powerplant weight.

(6) 240-kWe power delivered to ion engines.

The results from constant-acceleration data with asymptotic matching were as follows:

Parameter	Value
Propulsion time, days	~550
Specific impulse, s	4250
Initial acceleration, m/s^2	6.3×10^{-4}
Earth spiral time, days	~100
Coast time, days	~200
Jupiter spiral time, days	~50
Heliocentric time, days	~600
Propellant weight, lb	17,700

The initial thrust (coast plus earth spiral time) will require 300 days; the final thrust (coast plus Jupiter spiral time) will require 250 days. Therefore, the flight time (initial and final thrust times plus coast time) will be 750 days. The spacecraft will be required to orbit Jupiter for some time and send back environmental data to complete its mission. The time required to complete one orbit about Jupiter at 2×10^6 km is approximately 20 days. If a total of three orbits is assumed sufficient to complete

[3]Developed with the assistance of the JPL Systems Analysis Research Section.

the mission, an additional 60 days will be required. Therefore, the total emission time will be 810 days. The resulting mission profile is characterized in Fig. 13. The trajectory profile can be used to estimate effective sink temperatures, as well as requirements for protection against space radiation and meteoroids.

During the coast and orbit periods, the powerplant will have to maintain partial power to sustain the communications and hotel loads. It has been estimated that 10% electrical power will be sufficient to handle the hotel load. The thermal power required to maintain 10% electrical power will vary between 20 and 30%, depending upon the primary radiator temperature maintained. This will depend upon the spacecraft design and must be factored into the calculations for nuclear fuel burnup requirements.

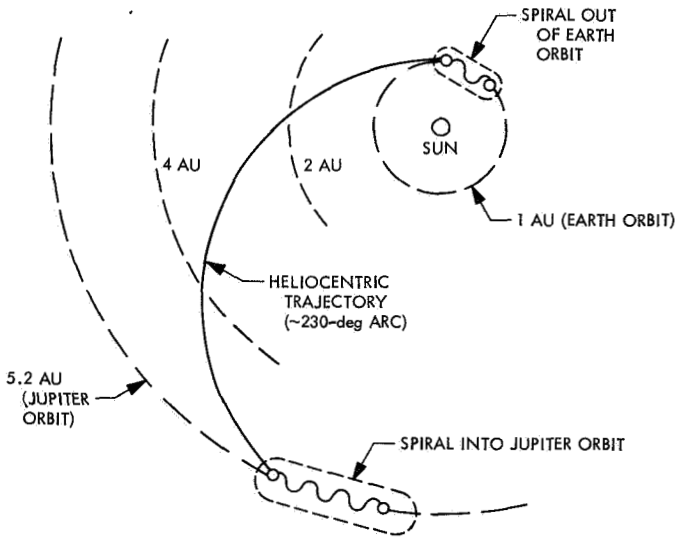


Fig. 13. Jupiter orbiter mission profile

The *Titan IIIC-7* fairing design, first-stage diameter, and required launch probability impose a volume constraint on the payload. With a 35-ft-long fairing, launch probability is high (99%). If a strengthened trans-stage is assumed, 10-ft-diam fairings up to 80 ft in length appear acceptable. It is presently estimated that a 70-ft-long fairing will be utilized for spacecraft containment within the nominal 10-ft diameter and for minimization of shielding requirements.

Launch-vibration characteristics are also an important constraint on spacecraft design. Peak acceleration spectral densities of 0.25 g^2/Hz are found in the region from 600 to 1000 Hz (Ref. 2).

References

1. Simon, N. K., and Sterns, J. W., Jr., "Evaluation of Nuclear-Electric Propulsion for Planetary Exploration by Graphic Analysis Techniques," Paper 68-545, presented at the AIAA Third Joint Propulsion Specialists Conference, Cleveland, Ohio, June 1968.

2. *Titan IIIC Payload Users Guide.* Martin-Marietta Corp., Denver, Colo., Jan. 1968.

XVII. Liquid Propulsion

PROPULSION DIVISION

A. Advanced Combustion Device Development,

R. W. Riebling

1. Introduction

The Liquid Propulsion Section has been conducting a number of research programs to evolve injector and thrust chamber design criteria applicable to the development of rocket engines for unmanned spacecraft, using the space-storable propellant combination oxygen difluoride (OF_2) and diborane (B_2H_6). The Advanced Combustion Device Development work unit has as one of its objectives the application of these criteria to the development of a prototype OF_2/B_2H_6 flight propulsion rocket engine. The prototype engines will be considered for use in the JPL space-storable propellants propulsion module, which is currently serving as a framework for technology advancement with these propellants. With a nominal vacuum thrust level of 1000 lbf, it will be capable of performing planetary missions around Mars, Venus or Jupiter. The adopted technical approach emphasizes the application of existing technology, rather than the generation of new technology, with the intent of exposing any problems encountered in an actual application. These will be solved as they arise, or made the subject of additional new technology programs, as appropriate. Another program objective is the demonstration of the basic soundness of currently accepted design concepts, in terms of performance and stability levels, long-duration durability, and inherent reliability.

Of the problem areas uncovered by the technology effort thus far, perhaps the most stringent is that of guaranteeing the structural integrity of an injector/chamber combination in very long duration firings under the severe environmental conditions imposed by the OF_2/B_2H_6 combustion process. This involves suppressing throat erosion, chamber streaking, and injector melting at flame temperatures on the order of 7000°F, in the presence of extremely reactive and corrosive products of combustion. The adiabatic wall temperature in the chamber and at the throat must be minimized without incurring undue performance penalties, and heat flux to the injector and chamber must be reduced. Materials are required that are resistant to mechanical erosion by liquid or possibly solid products in the combustion gas stream, and to chemical attack as well. Even at wall temperatures as low as 3500°F many materials are chemically attacked by the combustion products, which contain such corrosive species as fluorine, oxygen, and boron.

The thrust chamber design criteria evolved to date have been based on the concept of a passively cooled chamber, combined with an injector having a controlled gradient in mixture ratio. That is, the injector includes a central core region, containing 70 or 80% of the total mass flow and operating at or slightly above the design mixture ratio, and an outer or "barrier" zone, containing only 20 or 30% of the total mass flow and operating at

a much-reduced mixture ratio. The fuel-rich conditions prevailing in this barrier zone reduce both the adiabatic wall temperature and the concentrations of chemically undesirable species in the region immediately adjacent to the walls, thus helping to maintain the chamber's structural integrity. Nickel has emerged as the most durable injector material for use with fluorinated oxidizers such as OF_2, and carbonaceous or graphitic materials have been reported to be the most promising candidates for use as flame liners and throat inserts in chambers thermally insulated by ablative or other materials.

Accordingly, the technical approach adopted to develop a flight-prototype OF_2/B_2H_6 rocket engine has been based on a passively cooled chamber, with a carbonaceous or graphitic inner liner, and an injector made of nickel which incorporates some degree of barrier zone protection. In the event that the rocket engine must be "buried" within a spacecraft, its outside wall temperature and heat flux per unit area may not be permitted to exceed certain values. Therefore, an additional requirement imposed upon the engines currently being developed is that they be adequately thermally insulated.

Four competing versions of this general design concept are being evaluated. One was designed and fabricated at JPL, and the other three were made for JPL by industrial contractors. For reasons of economy, and also because there may ultimately be a need for an engine of this size, the initial evaluation is being conducted at the 200-lbf-thrust level. To assure commonality, all chambers were manufactured in compliance with a single JPL control drawing, and will be evaluated with interchangeable, JPL-furnished injectors. Based on performance and durability in 1000-s firing tests, the most promising chamber concept will be selected for final optimization and the development of prototype flight hardware at the 1000-lbf vacuum thrust level.

This article will describe the design and fabrication of one of the four passively cooled, insulated graphite thrust chamber assemblies. Future articles will describe the other chambers and the injectors, and will report the results of the experimental evaluations.

2. General Requirements

The basic ground rules and nominal operating conditions for which the chambers were designed and fabricated are shown in Table 1. The relatively large contraction area ratio is a result of the large injector diameter (2.6 in.) required to accommodate the number of injection elements necessary to assure an adequate mixture ratio gradient. Lower contraction ratios are conducive to improved secondary atomization, and hence somewhat higher combustion efficiencies with this propellant combination, and would be preferred for larger scale engines. The expected slight effect on vaporization efficiency was offset in the present case by the use of a somewhat longer than usual chamber characteristic length (L^*) of 20 in.

All the chambers have an expansion area ratio of 5.0 and are provided with an external lip (Fig. 1) to interface with a straight-tube diffuser at the test facility. This large expansion ratio (for ground level tests) was necessary to provide a large enough throat and divergent

Table 1. General requirements of 200 lbf OF_2/B_2H_6 rocket engines

Core mixture ratio	3.85
Nominal chamber pressure	100 psia
Total propellant flow rate	0.55 lbm/s
Throat diameter	1.162 in.
Chamber inside diameter	2.6 in.
Chamber characteristic length	20 in.
Contraction area ratio	5.0
Expansion area ratio	5.0
Duration	1000 s (continuous)
Restart capability	Multiple
Throttling capability	None required
Maximum allowable outside wall temperature	600°F

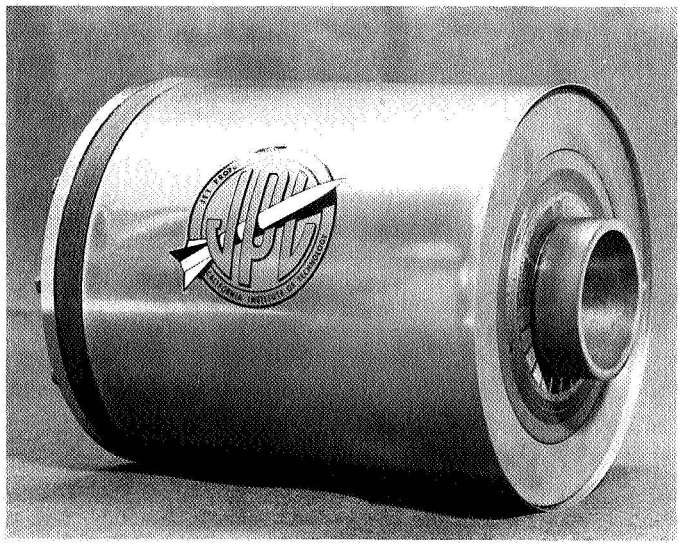


Fig. 1. Complete Radiabatic thrust chamber assembly

nozzle heat-transfer area to assure the meaningful evaluation of each chamber concept. In the absence of an altitude chamber, the diffuser will assure that these nozzles flow full during the firings.

The maximum allowable outside wall temperature shown in Table 1 is based on the projected requirements for a thrust chamber completely enclosed within a spacecraft.

3. Radiabatic Thrust Chamber Concept

One of the four thrust chamber concepts being evaluated is the Radiabatic[1] thrust chamber. A complete thrust chamber is depicted in Fig. 1; an exploded sectional view is depicted in Fig. 2; and its physical characteristics are summarized in Table 2.

The heart of the Radiabatic concept is the pyrolytic graphite heat conduction wedge assembly which surrounds the monolithic graphite chamber liner. Pyrolytic graphite is anisotropic; that is, in one plane (A–B plane), it is a very good heat conductor, while in a perpendicular direction, known as the C-direction, it behaves as an excellent thermal insulator. In the thrust chamber, the wedges are so oriented that their C-direction is circumferential. Therefore, a maximum amount of heat will be conducted along them in the radial and axial direction, while circumferential conduction is minimized. Most of the heat entering the liner in the throat region, where temperatures are a maximum, will thus be conducted axially to the relatively cooler head end of the chamber, and rejected to the combustion gas stream in the vicinity of the injector, and to the liquid-propellant-cooled injector itself. The Radiabatic concept is thus similar to other metallic or bimetallic conduction-cooled thrust chambers, with one significant distinction: in the Radiabatic chamber, heat is rejected only to the injector and to the relatively cool combustion gases in the barrier zone. No auxiliary film-cooling should be required, because the graphitic materials can operate at higher temperatures than most metals.

The pyrolytic graphite wedge assembly is surrounded by a layer of Grafoil[2] and a pyrolytic graphite insulator. The latter insulator has its A–B plane axially oriented, and its C-direction radially oriented, to minimize radial heat conduction. Its effect is to drop the wall temperature at its outer periphery to about 2000°F, so that the entire liner–insulator assembly can be contained within a columbium alloy shell, which is the main structural element of the thrust chamber. The final drop in temperature from columbium shell to the outer steel can is accomplished by two concentric layers of Dynaflex[3] insulation, which reduce the outside wall temperature to a predicted value of less than 600°F.

Table 2. Physical characteristics of the Radiabatic thrust chamber (less injector)

Overall length	10.25 in.
Maximum outside diameter	7.88 in.
Weight	31.25 lbm
External volume	458 in.^{3}

The JPL injector is mounted directly to the adapter flange shown in Fig. 2. Heat conduction into the propellant valve and the surrounding structure via this flange is minimized by the interposed silica–phenolic ablative insulator.

4. Fabrication and Assembly

The graphite chamber liners were machined from solid billets of commercially available, molded artificial graphite. One thrust chamber had a liner made of ATJS[4] grade graphite, while the second had a liner of regraphitized graphite (PT0248)[4]. The first material was chosen for its low cost and ready availability, and will serve as a reference material against which to compare more sophisticated products. The second was selected because of its reported superior thermal shock resistance. Further advantages of the monolithic liners include ease of fabrication and quality control because of one-piece construction from a single material. A chamber liner with the pyrolytic graphite retainer ring cemented in place is shown in Fig. 3.

The pyrolytic graphite heat conduction wedges were originally formed with rectangular axial sections. Twenty-four of these wedges were then laid up on a graphite mandrel with an equal number of Grafoil laminate separators, and bonded together with C7 Resin

[1]Manufactured by Thiokol Chemical Corp., Reaction Motors Div. Denville, N. J.

[2]Product of Carbon Products Div., Union Carbide Corp., New York, N. Y.

[3]Product of Johns Manville, New York, N. Y.

[4]Product of Carbon Products Div., Union Carbide Corp., New York, N. Y.

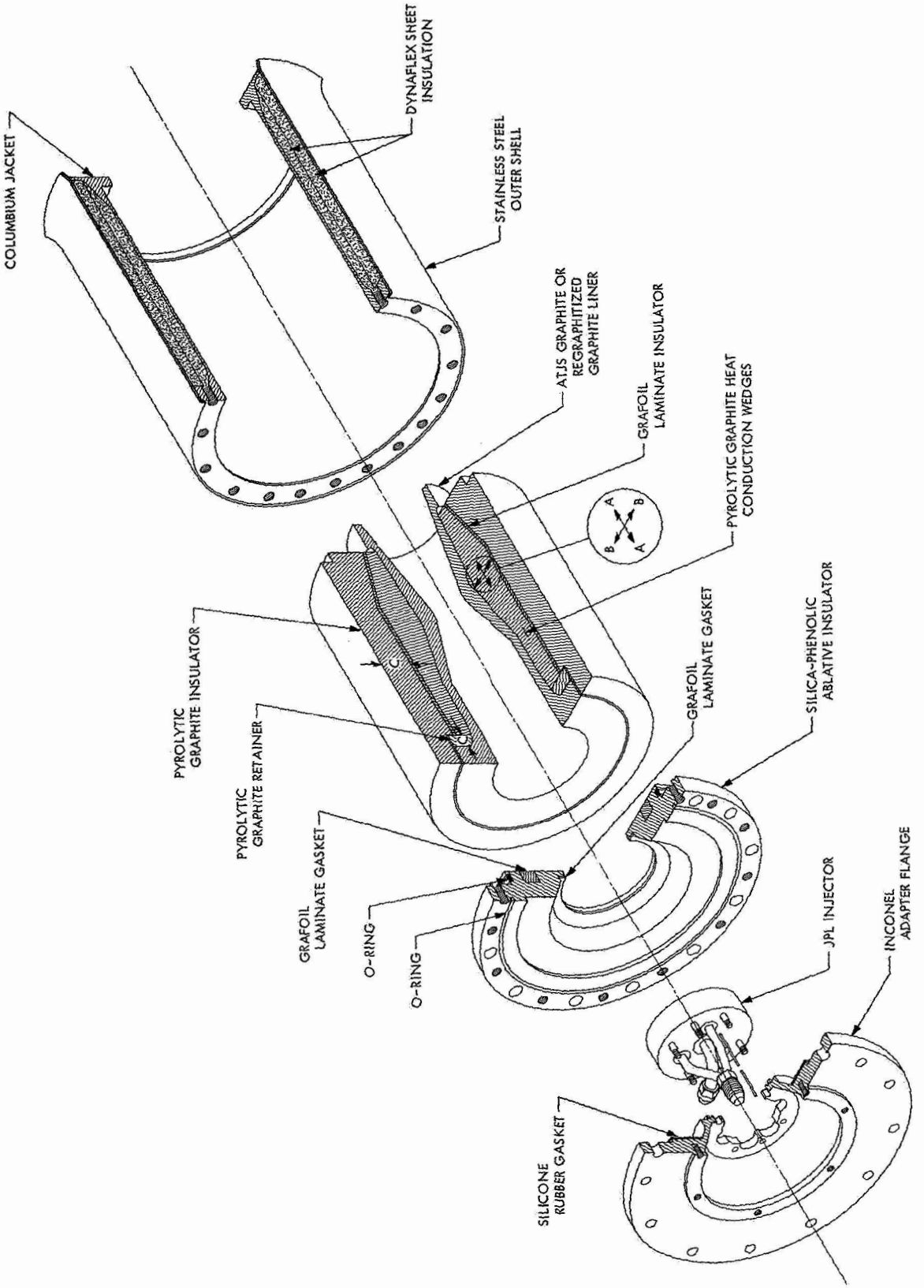


Fig. 2. Exploded sectional view of Radiabatic thrust chamber assembly

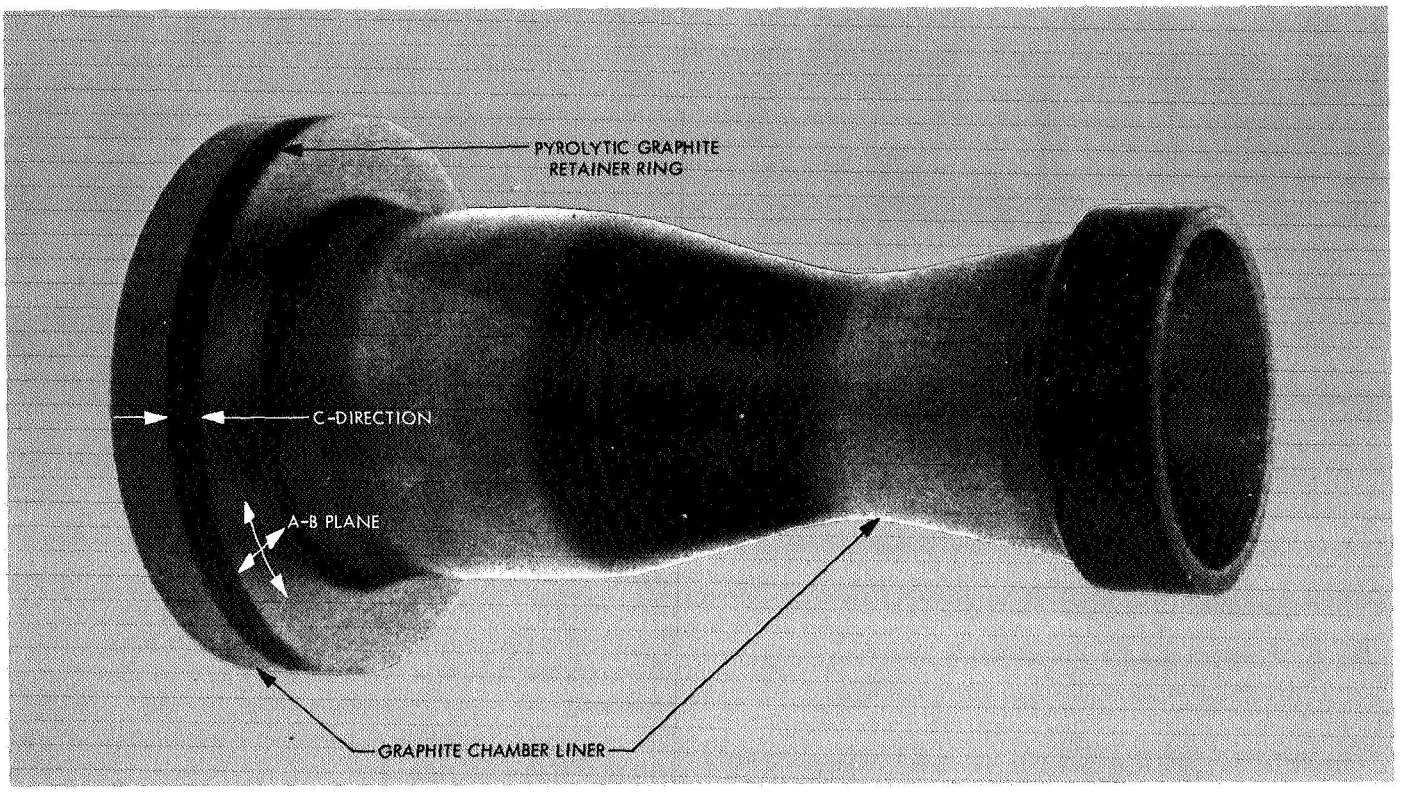


Fig. 3. Graphite chamber liner (photo courtesy of Thiokol Chemical Corp.)

(with Activator R)[5], as shown in Fig. 4. The separators further reduced circumferential heat conduction between wedges. Resin was intentionally omitted in one plane bisecting the cylinder longitudinally, however, so that the wedge assembly could be removed as a split cylinder. Subsequently, the internal and external contours shown in Fig. 2 were machined on a tracer lathe, the mandrel being destroyed in the process. The interior of the wedge assembly and the exterior of the chamber liner were lapped to assure intimate contact, and the split wedge assembly was cemented to the liner with C7 resin.

The pyrolytic graphite insulator was formed in one piece, with an internal taper at the aft end to match a corresponding external taper at the aft end of the wedge assembly (Fig. 2). A view of this part during the fabrication process is shown in Fig. 5, which also shows the additional material subsequently deposited. This insulator, together with an intervening layer of Grafoil laminate, was then cemented in place around the wedge assembly. The combined assemblies were finally cemented into the columbium alloy jacket assembly, two layers of Dynaflex insulation cut, shaped, and installed to surround the jacket, and the protective outer sleeve cemented in place.

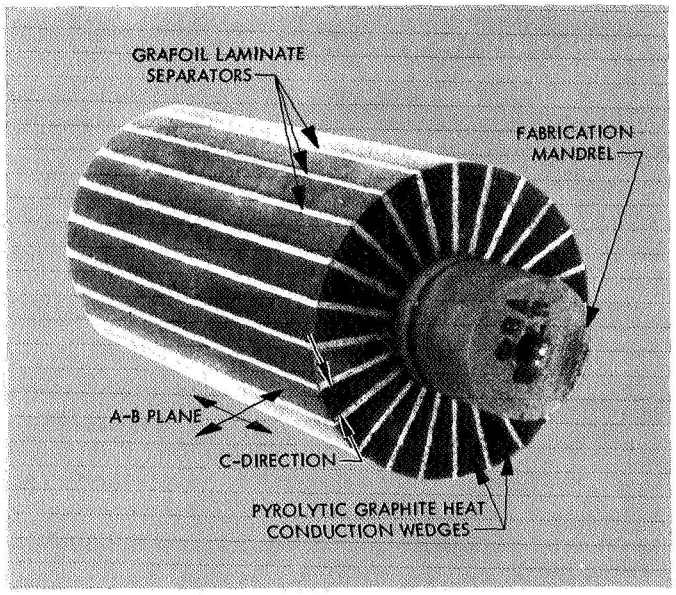


Fig. 4. Pyrolytic graphite wedge assembly before final machining (photo courtesy of Thiokol Chemical Corp.)

[5]Product of Armstrong Products Co., Warsaw, Ind.

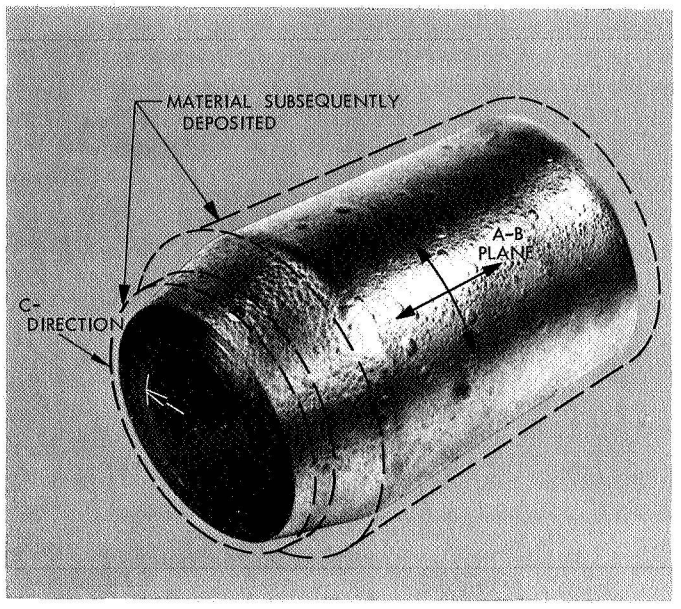


Fig. 5. Pyrolytic graphite insulator during fabrication (photo courtesy of Thiokol Chemical Corp.)

The ablative insulator and adapter flange were bolted to the columbium jacket assembly after installing the miscellaneous gaskets shown in Fig. 2.

B. Preliminary Burner Tests With Oxygen Difluoride and Diborane,

F. Gerbracht, R. Scott, and R. Kushida

Opposed-flow porous plate burner experiments were planned using diborane as fuel and oxygen difluoride (OF_2) as oxidizer. There have been no previous burner experiments reported on this system; hence, a number of unusual and potentially serious problems had to be resolved before a definite design could be evolved. A prototype opposed-flow porous plate burner was used; diborane was burned with air, and OF_2 was burned with ethylene.

A photograph of the test burner setup used in the OF_2–ethylene tests is shown in Fig. 6. The OF_2 was stored as a gas in the 500-ml stainless steel bottle shown at the right-hand side of the figure which also shows the pressure gauge, solenoid operated valve, the fittings containing a brass orifice plate, and the water-cooled burner head. The burner head contained a 0.5-in.-diam 0.125-in.-thick stainless steel porous disk[6] as the flame

[6]Manufactured by ASCO Sintering Company; it is Grade 1015 which filters particles larger than 15 μ.

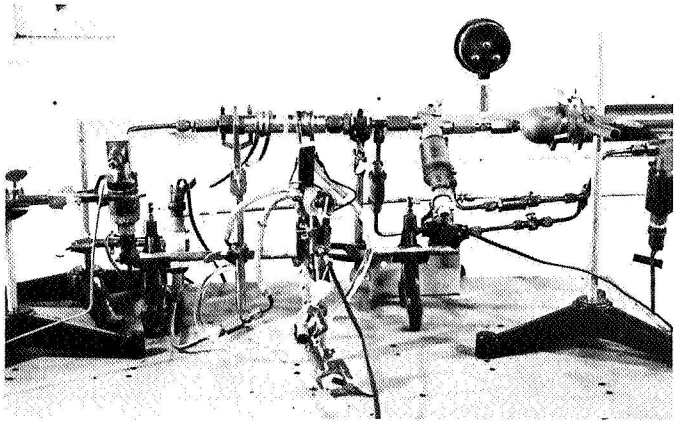


Fig. 6. Test burner set-up for OF_2–ethylene tests (opposed flow)

holder. The orifice served as a flowmeter. The 0.004-in.-diam orifice was used at low flow rates, and the 0.009 in. diam orifice was used when higher flow rates were desired. The ethylene fuel was fed through a porous disk burner of exactly the same dimensions as for the OF_2 side. It is shown in position in Fig. 6 so that an opposed-flow burner configuration is obtained.

The burner configuration for the diborane test was exactly similar to the OF_2 side of the burner shown in Fig. 6. The propellant tank was filled with diborane by connecting it to a supply of gaseous diborane and cooling the tank with liquid nitrogen. This froze the diborane in the propellant tank so that noncondensable gases could be pumped off. The tank was closed off and allowed to recover to room temperature.

During burning tests with gaseous diborane and air, thick deposits of a white congolomerated solid were found inside the diborane lines, around the metering orifice, and in back (i.e., upstream side) of the porous stainless steel burner. Further, on the air side of the porous plate, a thick solid deposit of white substance with some gray or blackish contaminant formed on the porous burner surface (Fig. 7(a)). Chemical analysis established that the deposits were boric oxide, with some metallic boron forming the dark contaminants. It was specifically determined that only minute and nontoxic amounts of boranes were present in these deposits. The deposits in the supply lines were eliminated by drying the lines and using a dry argon purge. The deposits on the porous plates could not be completely eliminated, but were reduced to an acceptable level by using higher gas flow rates through the porous plate. Figure 7(b)

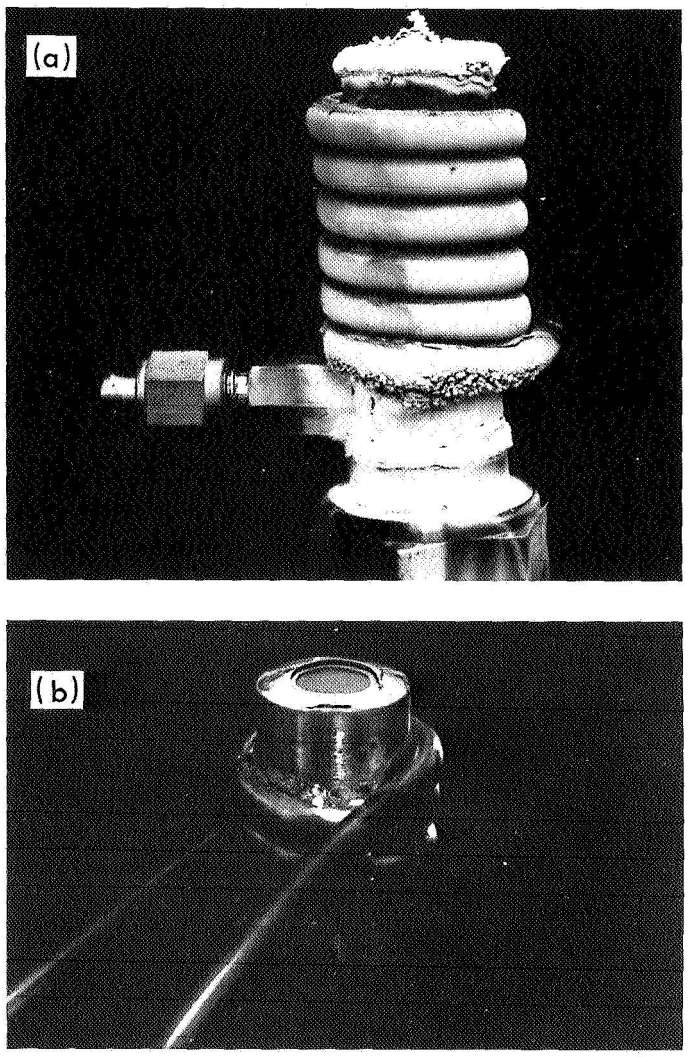


Fig. 7. Porous plate after test: (a) high gas flow rate of diborane (b) low gas flow rate of diborane

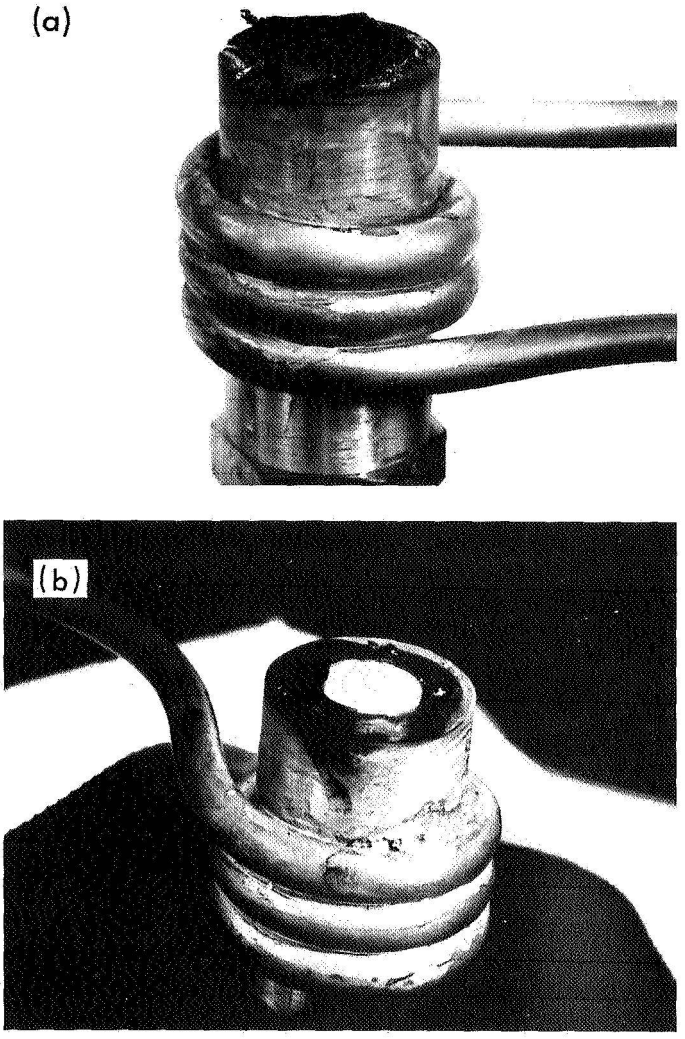


Fig. 8. Porous plate after test: (a) high gas flow rate of oxygen difluoride (b) low gas flow rate of oxygen difluoride

illustrates the appearance of the porous plate when a higher gas flow rate was used.

A tank and burner assembly, illustrated in Fig. 6, was passivated for 24 h at one atmosphere with OF_2. Ethylene was used as a fuel and injected through a porous plate burner head. The opposed flow configuration had the 0.5-in. diam stainless steel porous plates 0.25 in. apart.

At low flow rates the incomplete combustion of ethylene caused thick deposits of soot to appear on the fuel-side porous surface (Fig. 8(a)). Higher gas flow rates eliminated these deposits (Fig. 8(b)). In one run, the opposed flow gap was decreased to 0.080 in. in order to induce higher temperatures on the porous stainless surface. These were no visible signs of deterioration of the stainless steel porous plate due to exposure to OF_2 at an estimated 400 to 500°F for 3 min.

It was intended that the diborane be burned with OF_2 in the concluding series of tests. However, a leak was found in the diborane valve. Since the major questions concerning techniques for handling small quantities of OF_2 and diborane in the porous plate burner were answered, and since the replacement of the valve would require considerable reworking of the burner, these experiments were postponed until an improved system would be available.

These tests established that the opposed-flow burner could be used with OF_2 and with diborane without clog-

ging and without excessive oxidation of the stainless steel porous plate. The design and fabrication of the research version of the opposed-flow burner have been initiated. This burner will be operated with OF_2 and diborane at reduced pressure and with a gap spacing which will be variable during a run.

C. Porosity Surface Change of Shell 405 Catalyst, Preliminary Investigation, *A. Karbin*

1. Introduction

Ignition delays and high-pressure spiking were encountered in cold starts of hydrazine–Shell 405 engines when used in conjunction with a hydrazine gas plenum system. This system is a method of generating and storing gases for use as needed by attitude-control gas jets. Sharp high-pressure rises in the catalytic engine tend to disintegrate the catalyst, which could then contaminate the generated gases with particles. The one cause of pressure spikes has been hypothesized to be "catalyst poisoning" or temporary deactivation, which is alleged to result from gases such as ammonia, nitrogen, etc., entering the catalyst pores and becoming adsorbed on the active sites, thus preventing the hydrazine—liquid or vapors—from fully contacting the iridium surfaces. However, if the outside surface of the catalyst pellets (⅛-in.-diam and ⅛-in.-long cylinders or of finer mesh catalyst) becomes closed, then the cause of ignition delays may be due to a mechanical surface change rather than poisoning. Furthermore, the use of catalytically decomposed hydrazine appears to be increasing significantly; thus a deeper understanding of the causes of spiking is necessary.

2. Investigation of Surface Porosity

A preliminary investigation reveals a change in the surface porosity of Shell 405 ⅛-in. pellets with use. There appears to be a closing or glazing of the surface pores after hydrazine firings, which tends to decrease the exposed impregnated iridium surface and could be one of the causes of ignition delays. The resultant pressure spikes are one cause of swift catalyst attrition.

The surface pore characteristics are portrayed by the electron scan microscopic pictures of ⅛-in. pellets of Shell 405 with various histories. Figure 9 is of a ⅛-in. cylindrical pellet that has not been treated with iridium. Figure 10 shows a treated (Shell 405) pellet that has not been fired in an engine. Figure 10(a) is a photomicrograph of the polished cross section of the latter pellet. Fig. 11(a) to (d) are Electron microscope photos of a pellet that has been fired in a hydrazine engine for less than a total of 100 s in 5-s pulses spaced over a period of several weeks. The photomicrograph of Fig. 11(e) is of the cross section of the latter pellet.

There appears to be a similarity between all the cross-sectioned photomicrographs (Figures 10(a) and 11(e)), but there is a definite difference between the electron scan pictures taken of their respective outer surfaces. The catalyst without iridium (Fig. 9(c)) and the one impregnated with iridium but not fired (Fig. 10(b)) appear to have well-defined porous surfaces. In contrast, the pictures (Fig. 11(b) and (c)) of a short-pulse-fired engine show a lack of definitive porosity.

3. Conclusions

The pictures indicate that there is a greater change in porosity on the surface than in the internal structure. Thus, one may theorize that the closing of the surface pores may be one cause of ignition delays by not allowing the hydrazine liquid or gas to contact the total impregnated internal and external iridium pore surfaces.

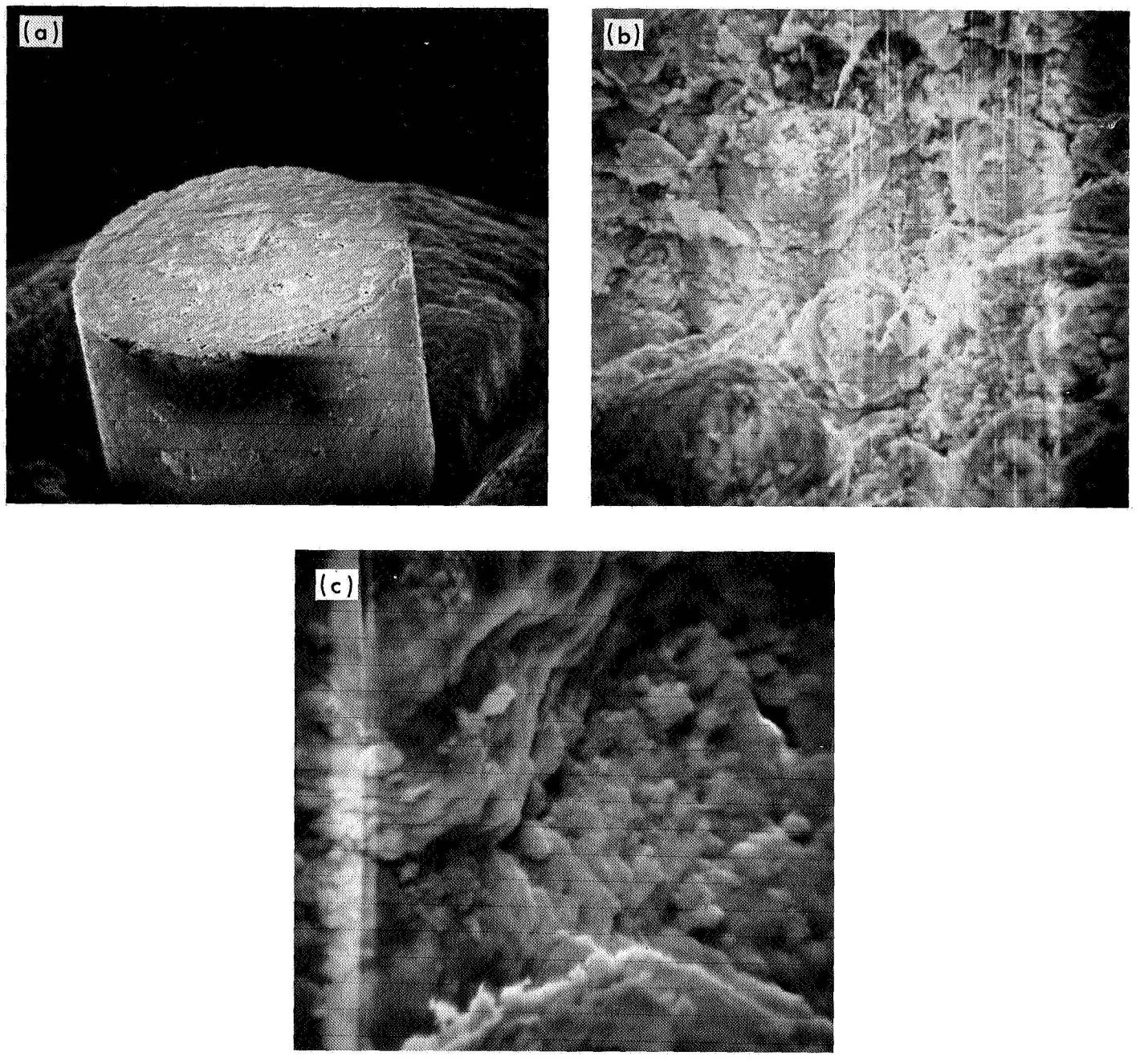


Fig. 9. *Catalyst prior to iridium treatment* (⅛ × ⅛-in. pellet, calcined at 700°C). Electron scan microscope at: (a) 18 X, (b) 900 X, (c) 4500 X

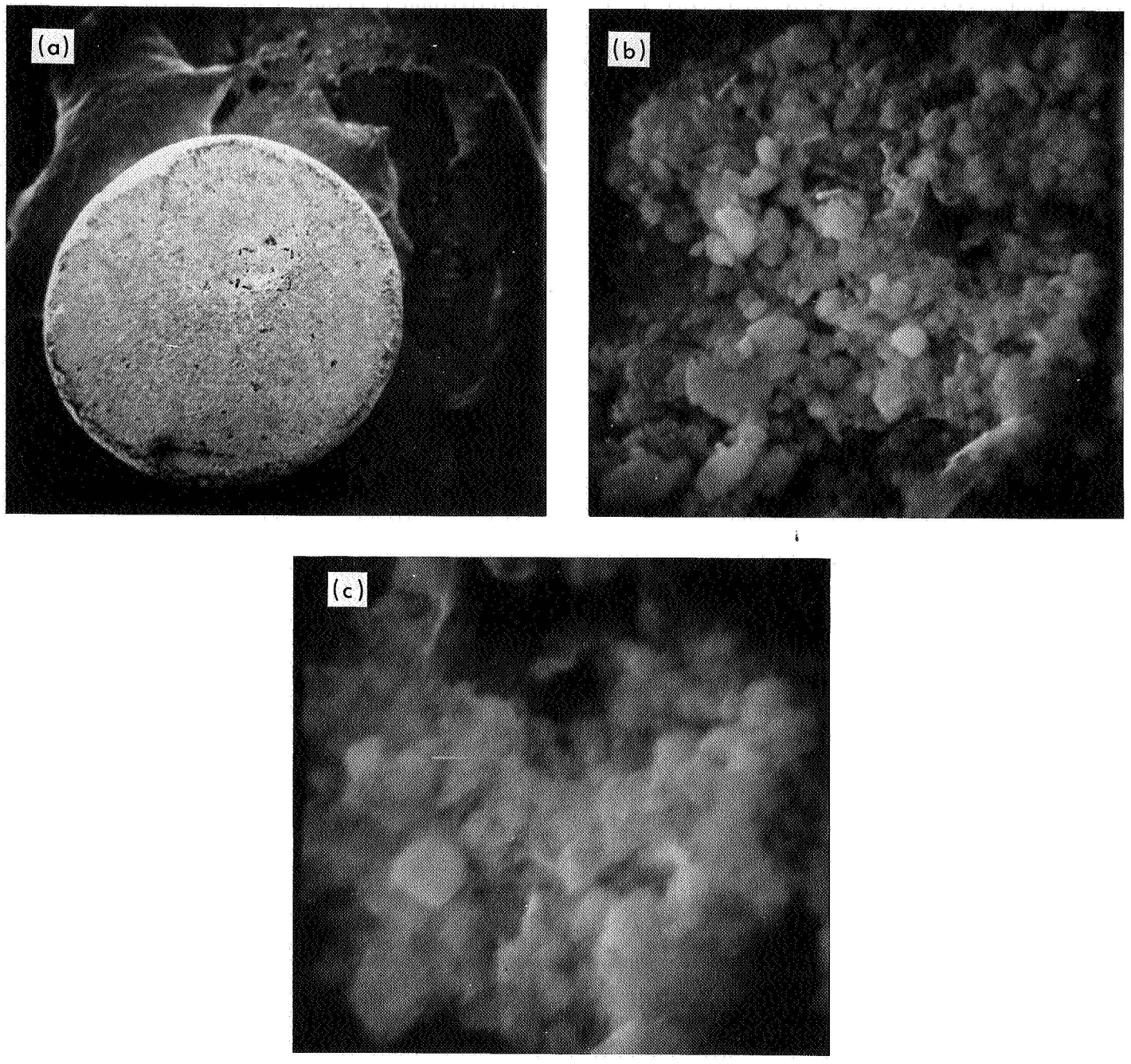
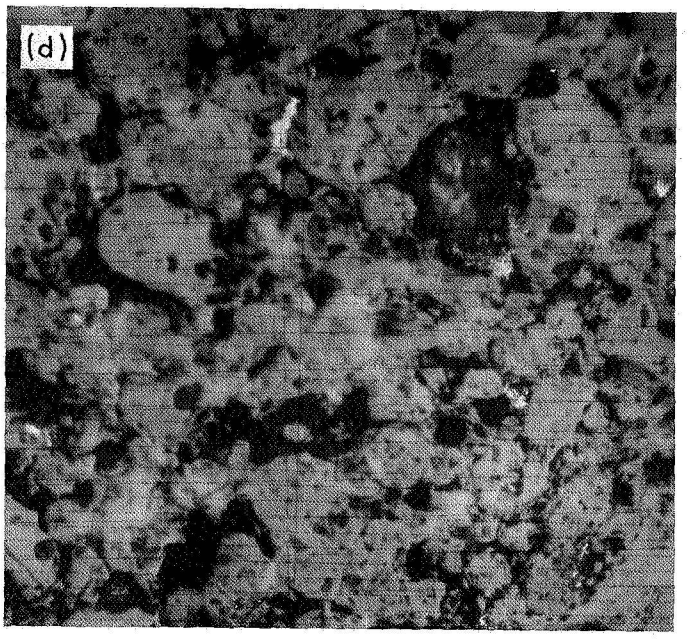
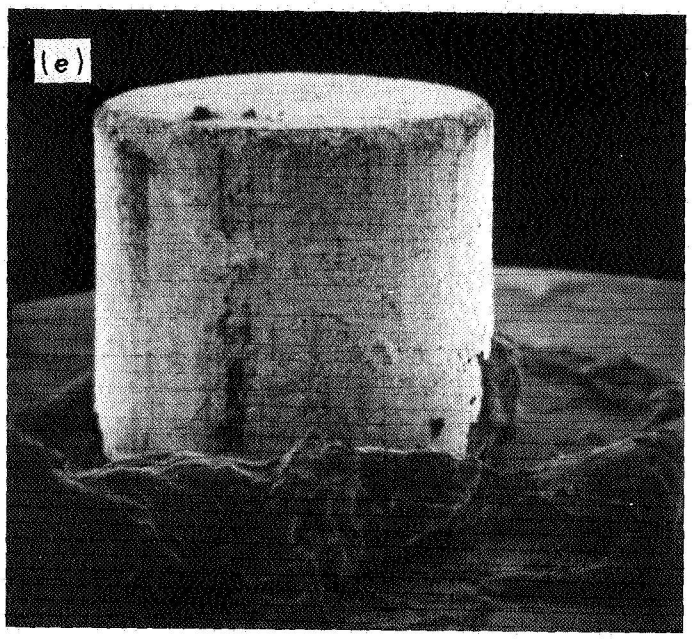


Fig. 10. Catalyst iridium treatment (Shell 405), ⅛ × ⅛-in. pellet, not fired. Electron scan microscope at: (a) 20 X, (b) 9900 X, (c) 19,800 X, (d) cross section at 200 X, (e) side view at 18 X, (f) portion of side view at 9350 X





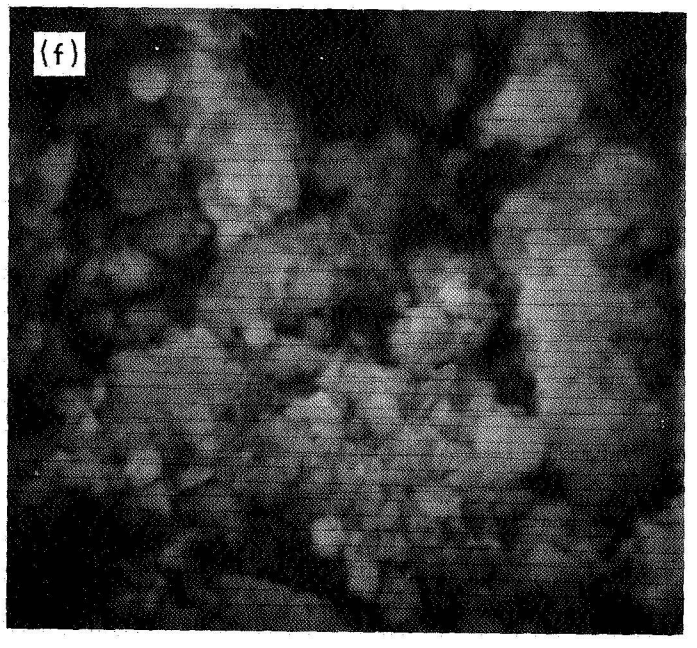


Fig. 10 (contd)

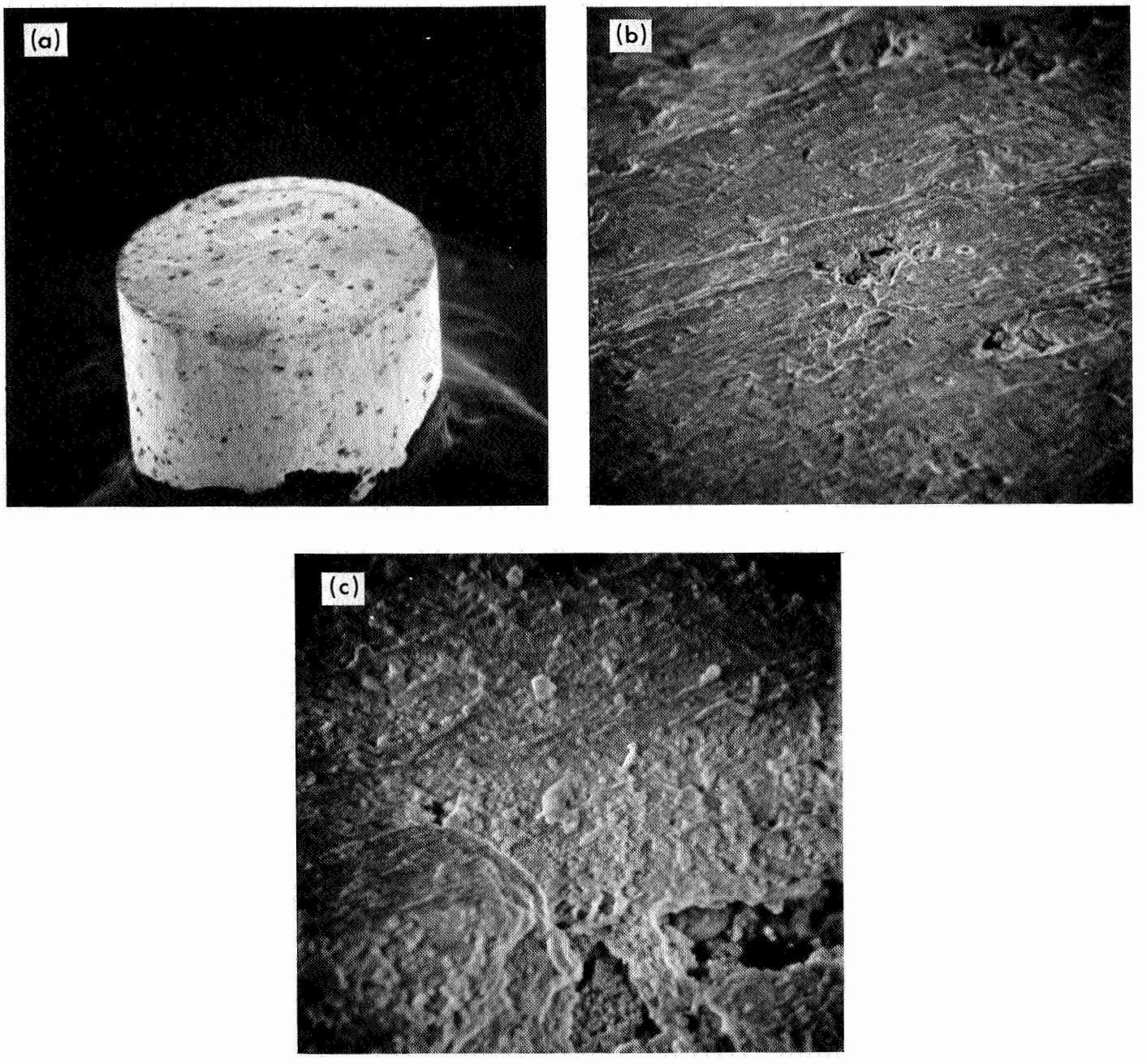


Fig. 11. Catalyst pellet (Shell 405) fired for less than 100 s in 5-s pulses over several weeks. Electron scan microscope at: (a) 18 X, (b) 1180 X, (c) 1800 X, (d) 17,500 X, (e) cross section at 200 X

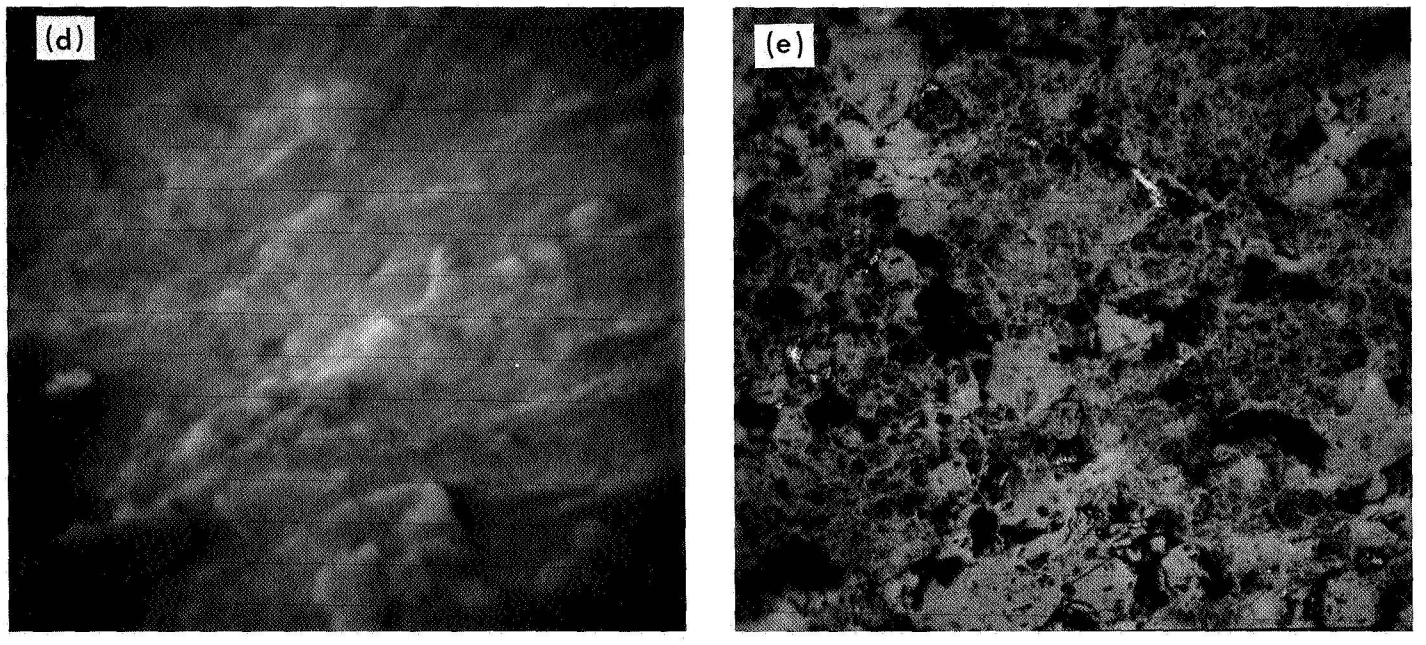


Fig. 11 (contd)

XVIII. Propulsion Systems Analysis and Advanced Engineering

PROPULSION DIVISION

A. Sterilization and Impact Testing of Experimental Pyrotechnic Electronic Unit,

M. L. Moore

1. Introduction

Most spacecraft pyrotechnic subsystems built by JPL have successfully employed a capacitive-discharge squib firing unit. The advantages of this unit—small demand on the spacecraft power subsystem, simplicity of design, immunity from false firing, and solid-state construction—make it attractive for use in pyrotechnic subsystems in hard-landing capsules.

The wet foil tantalum capacitors used for energy storage in all previous capacitive-discharge pyrotechnic subsystems successfully withstood a dry-heat sterilization test[1]; however, the results of impact tests were less than satisfactory.[2] An investigation was conducted to determine if a typical capacitive-discharge firing circuit and chassis fabricated with presently available electronic components would continue to operate normally after being subjected to sterilization and impact tests.

[1]Moore, M. L., *Thermal Sterilization Test Mariner C Pyrotechnic Control Unit,* Aug. 1, 1968 (JPL internal document).

[2]Bilodeau, K. R., and Evans, K. C., *High Impact Survival of Electronic Components,* Sept. 13, 1968 (JPL internal document).

2. Description of Unit

The prototype firing unit used in the investigation was small (approximately 2.75 × 1.0 × 0.8 in.) and lightweight (3.11 oz), which would allow wide flexibility in pyrotechnic subsystem design. The unit before potting and conformal coating is shown in Fig. 1a. It was designed to deliver a firing pulse of at least 5 A, for a minimum time of 1 ms, to a 1-A, 1-W squib bridgewire with a capacitor bank potential of 50 V. A silicon-controlled rectifier switch, gated by a unijunction transistor, delivers the stored energy to the squib load.

The capacitors were expected to be the most susceptible of all components to failure or damage during the sterilization cycles and high-impact shocks. The Sprague, Type 137D capacitors selected for energy storage are hermetically sealed, sintered-anode, tantalum capacitors containing a jelled electrolyte. Their operating temperature limit is 175°C; working voltage is 75 V at 85°C; and capacitance is 66 μF $\pm$20%. Four of the capacitors were paralleled for energy storage.

All components were solid-state. With the exception of the silicon-controlled rectifier (a General Electric, high-reliability type C11DR1200), all components were commercial-grade and had undergone no prefabrication

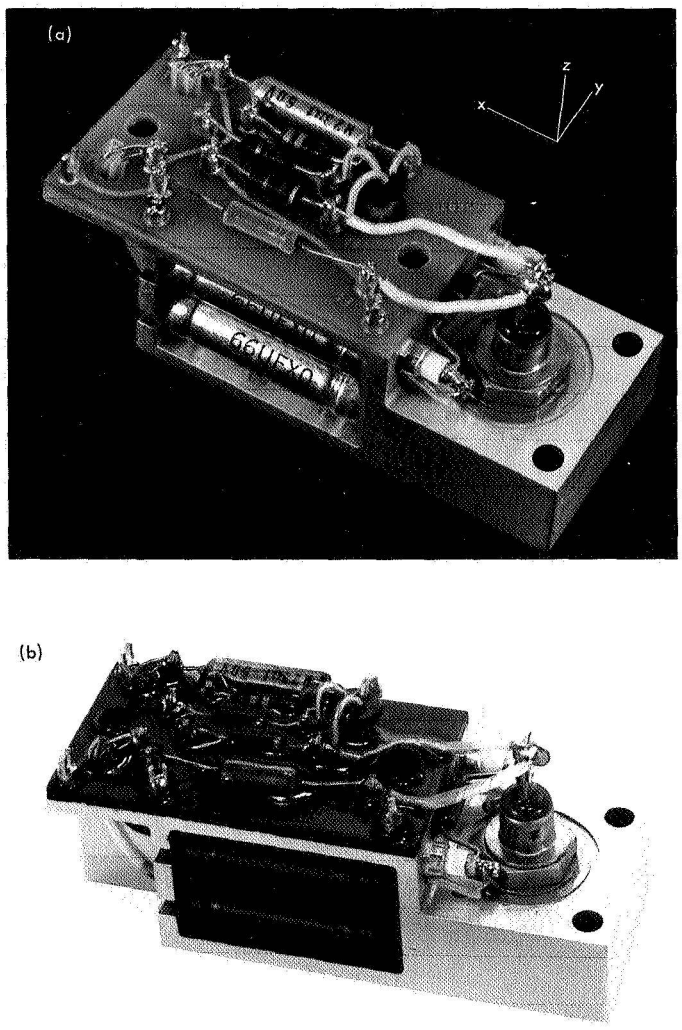


Fig. 1. Squib firing unit: (a) before potting and conformal coating, (b) after sterilization tests

burn-in or screening tests. The chassis was fabricated from 6061-T6 aluminum alloy. Total measured capacitance of the energy bank was 280 μF; measured leakage current of the capacitors totaled 0.85 μA at 50-V excitation. The maximum dissipation factor of any of the four capacitors was 3.5%.

3. Sterilization Tests

The unit was subjected to six, 36-h, dry-heat sterilization cycles at a temperature of 135°C. After its return to room temperature following each cycle, the unit was energized and discharged into a load that simulated a typical spacecraft squib circuit consisting of a 1-Ω, 1%, instrumentation resistor; a 1-Ω resistor simulating conductor resistance; and a 1-Ω resistor simulating squib resistance. Photographs of the current delivered to the load circuit were taken from an oscilloscope connected across the instrumentation resistor. Peak current output was 16 A; current at 1 ms was 5 A.

The most significant (but expected) difference in the unit after the sterilization cycles was the dark amber color of the Solithane conformal coating and potting compound, which was originally colorless. The affected areas can be seen in Fig. 1b. No other effects were observed visually or operationally.

4. Impact Tests

Following the sterilization tests, the unit was subjected to five impacts, beginning at 1000 *g* and increasing by 1000-*g* increments, in each of three mutually perpendicular axes (Fig. 1a) on the JPL "slingshot" shock tester. The shocks lasted 0.5–0.75 ms. After each impact, an operational test was made using the same instrumentation and test configuration as used during the sterilization test. Oscilloscope photographs of the delivered current were taken before and after the 15 impacts. No change was observed visually or operationally.

Final operational tests were made by firing 1-A, 1-W live squibs at an ambient temperature of 73°F. The circuit configuration for these firings was the same as that used after the sterilization cycles, except that the 2-Ω load resistor was replaced by a live squib. Photographs of oscilloscope current traces were taken during the squib firings.

5. Results

The firing unit met all design requirements after being subjected to dry-heat sterilization and shocks up to 5000 *g*. The only reduction in delivered firing current was observed after completion of the first sterilization cycle and was almost undetectable. The measured capacitance of the energy storage bank after conclusion of the tests was the same as that measured before the tests.

In the live squib firings, average currents of approximately 18.5 A were delivered until bridgewire burnout (at 0.16 and 0.218 ms), indicating that more-than-sufficient energy is available to fire this type of squib. Since the 1-Ω, 1% series resistor is included for instrumentation purposes only in a single squib firing circuit, it could be removed or replaced by a one-turn primary current transformer to allow additional energy to be delivered.